GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:54:38 ; Search time 4201.73 Seconds
(without alignments)
16071.974 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1188	100.0	1188	6	E63071	E63071 DNA encodin
2	1188	100.0	1188	15	AF021244	AF021244 Arabidops
3	1175	98.9	1194	15	AY099815	AY099815 Arabidops
4	1128	94.9	1139	15	BT000309	BT000309 Arabidops
5	1115.4	93.9	1157	15	AY089091	AY089091 Arabidops
6	1012	85.2	90341	15	AC024609	AC024609 Arabidops
7	1012	85.2	119942	15	AC007797	AC007797 Arabidops
8	975	82.1	975	6	AX412268	AX412268 Sequence
9	975	82.1	975	6	AX412469	AX412469 Sequence
10	975	82.1	975	6	AX505463	AX505463 Sequence
11	754	63.5	1115	15	AF337544	AF337544 Brassica
12	395.8	33.3	453	10	BX510604	BX510604 Arabidops
13	239.4	20.2	987	6	AX411601	AX411601 Sequence
14	202.6	17.1	1174	6	AX411607	AX411607 Sequence
15	199	16.8	1242	6	AX411617	AX411617 Sequence
16	199	16.8	1242	15	BT009214	BT009214 Triticum
17	194.4	16.4	1104	6	AX411609	AX411609 Sequence
18	188	15.8	1125	6	AX411611	AX411611 Sequence
19	178.8	15.1	1444	6	AX411613	AX411613 Sequence
20	176.8	14.9	1157	15	AY292526	AY292526 Ginkgo bi
21	169.8	14.3	1216	6	AX411603	AX411603 Sequence
22	169.4	14.3	1104	15	AF337545	AF337545 Brassica
23	169.2	14.2	992	15	AB056126	AB056126 Citrus un
24	167.6	14.1	1092	15	AF160869	AF160869 Citrus si
25	163.4	13.8	1135	6	E63072	E63072 DNA encodin
26	163.4	13.8	1135	15	AF134302	AF134302 Arabidops
27	161.8	13.6	1025	15	BT002898	BT002898 Arabidops
28	151.6	12.8	1340	6	E63070	E63070 DNA encodin
29	151.6	12.8	1340	15	AB025025	AB025025 Chenopodi
30	151.6	12.8	1340	15	AF134301	AF134301 Chenopodi
31	146.6	12.3	1159	15	BT013523	BT013523 Lycopersi
32	139.8	11.8	128855	14	AP007972	AP007972 Lotus cor
33	114.4	9.6	87286	15	AB026651	AB026651 Arabidops
34	110.8	9.3	1156	6	AX411625	AX411625 Sequence
35	110.8	9.3	1156	15	BT008923	BT008923 Triticum
36	109.2	9.2	1302	6	AX411605	AX411605 Sequence
37	98.6	8.3	916	6	AX411615	AX411615 Sequence
38	98.6	8.3	916	15	BT009115	BT009115 Triticum
39	82.6	7.0	1274	6	AX411627	AX411627 Sequence
40	73	6.1	884	15	AF337546	AF337546 Brassica
41	67	5.6	1300	6	E63080	E63080 DNA encodin
42	56.8	4.8	242	15	ATH524543	AJ524543 Arabidops
c 43	56.6	4.8	110000	15	AP008208_107	Continuation (108
c 44	56.6	4.8	169378	15	AP004001	AP004001 Oryza sat
c 45	53	4.5	110000	15	AP008216_142	Continuation (143

ALIGNMENTS

RESULT 1
E63071
LOCUS E63071 1188 bp DNA linear PAT 27-AUG-2002
DEFINITION DNA encoding chlorophylase and plant transformed by it.
ACCESSION E63071
VERSION E63071.1 GI:22553554
KEYWORDS JP 2001086990-A/2.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1188)
AUTHORS Tsuchiya,T., Ota,H., Takamiya,K., Harada,S. and Nakat,K.
TITLE DNA encoding chlorophylase and plant transformed by it
JOURNAL Patent: JP 2001086990-A 2 03-APR-2001;
KAGOME CO LTD
COMMENT OS Arabidopsis thaliana (thale cress)
PN JP 2001086990-A/2
PD 03-APR-2001
PF 20-SEP-1999 JP 1999266181
PI TORU TSUCHIYA,HIROYUKI OTA,KENICHIRO TAKAMIYA,SATOSHI HARADA,
PI KENGO NAKATA
PC C12N15/09,A01H5/00,C12N5/10,C12N9/16//(C12N9/16,C12R1:19), PC C12N15/00,
PC C12N5/00
CC
FH Key Location/Qualifiers
FT CDS (49). .(1020).
FEATURES Location/Qualifiers
source 1. .1188
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
ORIGIN

Query Match 100.0%; Score 1188; DB 6; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60

```
Db          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60

Qy         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120

Qy        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180

Qy        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240

Qy        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300

Qy        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360

Qy        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420

Qy        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480

Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540

Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600

Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660

Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720

Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780

Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840

Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900

Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960

Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020

Qy       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080

Qy       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140

Qy       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188


RESULT 2
AF021244
LOCUS       AF021244               1188 bp    mRNA    linear   PLN 17-APR-1998
DEFINITION  Arabidopsis thaliana coronatine-induced protein 1 (CORI1) mRNA,
            complete cds.
ACCESSION   AF021244
VERSION     AF021244.1  GI:2460202
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1188)
  AUTHORS   Benedetti,C.E., Costa,C.L., Turcinelli,S.R. and Arruda,P.
  TITLE     Differential expression of a novel gene in response to coronatine,
            methyl jasmonate, and wounding in the Coi1 mutant of Arabidopsis
  JOURNAL   Plant Physiol. 116 (3), 1037-1042 (1998)
   PUBMED   9501136
REFERENCE   2  (bases 1 to 1188)
  AUTHORS   Benedetti,C.E., Costa,C.L., Turcinelli,S.R. and Arruda,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-AUG-1997) Centro de Biologia Molecular e Engenharia
            Genetica, University of Campinas, Campinas, SP 13083970, Brazil
FEATURES             Location/Qualifiers
     source          1. .1188
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia; Landsberg erecta"
                     /db_xref="taxon:3702"
                     /tissue_type="seedlings; leaves; flowers; siliques"
     gene            1. .1188
                     /gene="CORI1"
                     /note="ATHCOR1; coronatine-induced gene; induced in
                     response to the phytotoxin coronatine, methyl jasmonate
                     and wounding"
     CDS             49. .1023
                     /gene="CORI1"
                     /note="similarity to certain plant and animal hydrolases;
                     includes a P-loop, a potential ATP/GTP binding site"
                     /codon_start=1
                     /product="coronatine-induced protein 1"
                     /protein_id="AAC13947.1"
                     /db_xref="GI:2460203"
                     /translation="MAAIEDSPTFSSVVTPAAFEIGSLPTTEIPVDPVENDSTAPPKP
                     VRITCPTVAGTYPVVLFFHGFYLRNYFYSDVLNHIASHGYILVAPQLCKLLPPGGQVE
                     VDDAGSVINWASENLKAHLPTSVNANGKYTSLVGHSRGGKTAFAVALGHAATLDPSIT
                     FSALIGIDPVAGTNKYIRTDPHILTYKPESFELDIPVAVVGTGLGPKWNNVMPPCAPT
                     DLNHEEFYKECKATKAHFVAADYGHMDMLDDDLPGFVGFMAGCMCKNGQRKKSEMRSF
                     VGGIVVAFLKYSLWGEKAEIRLIVKDPSVSPAKLDPSPELEEASGIFV"
ORIGIN

  Query Match           100.0%;  Score 1188;  DB 15;  Length 1188;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60

Qy         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120

Qy        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180

Qy        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240

Qy        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300

Qy        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360

Qy        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420

Qy        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480

Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540

Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600

Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660

Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720

Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780

Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840

Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900

Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960

Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020

Qy       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080

Qy       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140

Qy       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188

RESULT 3
AY099815
```

```
LOCUS       AY099815               1194 bp    mRNA    linear   PLN 06-MAY-2002
DEFINITION  Arabidopsis thaliana unknown protein (At1g19670) mRNA, complete
            cds.
ACCESSION   AY099815
VERSION     AY099815.1  GI:20466697
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1194)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2002) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G.,  Lam,B., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M, (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford)  contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1194
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL09-95-F15 (R25648)"
                     /ecotype="Columbia"
                     /note="This clone is in pBluescript"
     gene            1. .1194
                     /gene="At1g19670"
                     /note="synonym: F14P1.11"
     CDS             48. .1022
                     /gene="At1g19670"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAM20666.1"
                     /db_xref="GI:20466698"
                     /translation="MAAIEDSPTFSSVVTPAAFEIGSLPTTEIPVDPVENDSTAPPKP
                     VRITCPTVAGTYPVVLFFHGFYLRNYFYSDVLNHIASHGYILVAPQLCKLLPPGGQVE
                     VDDAGSVINWASENLKAHLPTSVNANGKYTSLVGHSRGGKTAFAVALGHAATLDPSIT
                     FSALIGIDPVAGTNKYIRTDPHILTYKPESFELDIPVAVVGTGLGPKWNNVMPPCAPT
                     DLNHEEFYKECKATKAHFVAADYGHMDMLDDDLPGFVGFMAGCMCKNGQRKKSEMRSF
                     VGGIVVAFLKYSLWGEKAEIRLIVKDPSVSPAKLDPSPELEEASGIFV"
ORIGIN

  Query Match           98.9%;  Score 1175;  DB 15;  Length 1194;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1175;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 60
```

```
Qy        62 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 120

Qy       122 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 180

Qy       182 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 240

Qy       242 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 300

Qy       302 TTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 360

Qy       362 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 420

Qy       422 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 480

Qy       482 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 540

Qy       542 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 600

Qy       602 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 660

Qy       662 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 720

Qy       722 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 780

Qy       782 ATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGA 841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 ATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGA 840

Qy       842 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 900

Qy       902 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 960

Qy       962 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1021
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1020

Qy      1022 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1080

Qy      1082 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1140

Qy      1142 TCGATCAAAACATTTGTTATAGTTTTACCCCAAAA 1176
             |||||||||||||||||||||||||||||||||||
Db      1141 TCGATCAAAACATTTGTTATAGTTTTACCCCAAAA 1175
```

RESULT 4
BT000309

```
LOCUS       BT000309                1139 bp    mRNA    linear   PLN 19-SEP-2002
DEFINITION  Arabidopsis thaliana unknown protein (At1g19670) mRNA, complete
            cds.
ACCESSION   BT000309
VERSION     BT000309.1  GI:23198201
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1139)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2002) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H.,
            Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,
            Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M.,
            Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S.,
            Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford) contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1139
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="U25648"
                     /ecotype="Columbia"
                     /note="This clone is in pENTR/SD-dTopo   This is a cloned
                     PCR product using RIKEN clone RAFL09-95-F15 (AY099815) as
                     a template"
     gene            1. .1139
                     /gene="At1g19670"
                     /note="synonym: F14P1.11"
     CDS             1. .975
                     /gene="At1g19670"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAN15628.1"
                     /db_xref="GI:23198202"
                     /translation="MAAIEDSPTFSSVVTPAAFEIGSLPTTEIPVDPVENDSTAPPKP
                     VRITCPTVAGTYPVVLFFHGFYLRNYFYSDVLNHIASHGYILVAPQLCKLLPPGGQVE
                     VDDAGSVINWASENLKAHLPTSVNANGKYTSLVGHSRGGKTAFAVALGHAATLDPSIT
                     FSALIGIDPVAGTNKYIRTDPHILTYKPESFELDIPVAVVGTGLGPKWNNVMPPCAPT
                     DLNHEEFYKECKATKAHFVAADYGHMDMLDDDLPGFVGFMAGCMCKNGQRKKSEMRSF
```

VGGIVVAFLKYSLWGEKAEIRLIVKDPSVSPAKLDPSPELEEASGIFV"
ORIGIN

Query Match 94.9%; Score 1128; DB 15; Length 1139;
Best Local Similarity 100.0%; Pred. No. 1.6e-306;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy       109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy       169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy       229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy       289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy       349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy       409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy       469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy       529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy       589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy       649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy       709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy       769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy       829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy       889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy       949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy      1009 GGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAA 1068
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        961 GGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAA 1020

Qy       1069 AACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGT 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGT 1080

Qy       1129 ATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAA 1176
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAA 1128
```

RESULT 5
AY089091
LOCUS AY089091 1157 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 31589 mRNA, complete sequence.
ACCESSION AY089091
VERSION AY089091.1 GI:21407865
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
PUBMED 12093376
REFERENCE 2 (bases 1 to 1157)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Full-Length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1157)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or LAer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.
FEATURES Location/Qualifiers
source 1. .1157
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="31589"
ORIGIN

Query Match 93.9%; Score 1115.4; DB 15; Length 1157;
Best Local Similarity 99.5%; Pred. No. 5.8e-303;
Matches 1151; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

```
Qy         2 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 61
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 64
```

```
Qy        62 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        65 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 124

Qy       122 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 181
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       125 CGACAACCGAGATACCTGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 184

Qy       182 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 241
             |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db       185 TGAGAATCACCTGTCCA--AGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 242

Qy       242 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       243 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 302

Qy       302 TTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 361
             ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db       303 TTCTTGTAGCCC--CAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 360

Qy       362 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 420

Qy       422 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 480

Qy       482 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 540

Qy       542 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 600

Qy       602 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 660

Qy       662 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 720

Qy       722 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 780

Qy       782 ATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGA 841
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       781 ATATGTTGGACGATGATTTGCCCGGTTTTGTT-GGTTTATGGCCGGTTGTATGTGTAAGA 839

Qy       842 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       840 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 899

Qy       902 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       900 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 959

Qy       962 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1021
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       960 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1019

Qy      1022 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1020 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1079

Qy      1082 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1080 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1139

Qy      1142 TCGATCAAAACATTTGT 1158
```

```
             ||||||||||||||||
Db      1140 TCGATCAAAACATTTGT 1156
```

RESULT 6
AC024609
LOCUS AC024609 90341 bp DNA linear PLN 30-OCT-2002
DEFINITION Arabidopsis thaliana chromosome I BAC F14P1 genomic sequence, complete sequence.
ACCESSION AC024609
VERSION AC024609.2 GI:7212002
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 90341)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Miranda,M., Brooks,S., Buehler,E., Chao,Q., Chin,C., Chiou,J., Choi,E., Gonzalez,A., Howng,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M., Lenz,C., Liu,A., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Shinn,P., Toriumi,M., Vaysberg,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 90341)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E., Chin,C., Chiou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 3 (bases 1 to 90341)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E., Chin,C., Chiou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 4 (bases 1 to 90341)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 5 (bases 1 to 90341)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT On Mar 9, 2000 this sequence version replaced gi:7121532.
Bases 1-59,676 of IGF clone F14P1 overlap with IGF clone F6F9, gb|AC007797.
e-mail for correspondence: arab@sequence.stanford.edu
Genes with similarity to proteins in the databases are named 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as

```
          'unknown proteins'.  Genes that are annotated based only on gene
          prediction software are described as 'hypothetical proteins'.  The
          gene prediction programs used to predict genes include:  Grail
          (Informatics Group, Oak Ridge National Laboratory,
          http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
          http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
          (S.M. Hebsgaard, et al., CBS, Technical University of
          Denmark,http://www.cbs.dtu.dk/NetPlantGene.html).
FEATURES             Location/Qualifiers
     source          1. .90341
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="I"
                     /clone="F14P1"
     misc_feature    1. .59676
                     /note="overlap with IGF clone F6F9, gb|AC007797, see
                     GenBank record for BAC clone F6F9 for annotation in this
                     region."
     gene            60210. .61960
                     /gene="F14P1.1"
     CDS             join(60210. .60507,60608. .60778,60878. .60927,61025. .61081,
                     61184. .61249,61387. .61461,61599. .61717,61807. .61960)
                     /gene="F14P1.1"
                     /note="Unknown protein"
                     /codon_start=1
                     /protein_id="AAF98407.1"
                     /db_xref="GI:9795589"
                     /translation="MRSLQAPVVCPSVRPRQLGVSALLVNCSVSKTRSLRKQFWGNQT
                     KNDKSQAATVNLRLHLRRYKSIKCLFSSHSDGTGSTAENFNENDEDYVKSSVLEAVEV
                     KSGPDGFMVKMKDGRQLRCVHNNPQGGNLPNYAPHSAIVLKMEDGTGLLLPIIVLEMP
                     SVLLMAAMTNVQIARPTMYQVVKDMVDKMGYEVRLVRVTTRVHEAYFAELYLSKVGDK
                     SDCVSFDLRPSDAINIAVRCKVPIQVNKYLAYSDGMRVIDSGKLSKQTPASDGLLFTE
                     LDRPNGQPCFDTKEFDLVRNMMQAVDEERYDEAGKLSLSHPPYLVVILVLINY"
     gene            63540. .65756
                     /gene="F14P1.2"
     CDS             join(63540. .63596,63685. .63951,64046. .64123,64205. .64306,
                     64387. .64436,64527. .64626,64741. .64833,64927. .65040,
                     65139. .65253,65359. .65756)
                     /gene="F14P1.2"
                     /note="Hypothetical protein"
                     /codon_start=1
                     /protein_id="AAF98408.1"
                     /db_xref="GI:9795590"
                     /translation="MWTNMIQWRRDFGTDTILEDFEFPELDEVLRYYPQGYHGVDKEG
                     RPVYIERLGKVDASKLMQVTTLERYLRYHVKEFEKTITVKFPACCIAAKRHIDSSTTI
                     LDVQGLGLKNFTKTARDLIIQLQKIDSDNYPETLHRMFIINAGSGFKLLWGTVKSFLD
                     PKTVSKIHVLGNKYQNKLLEMIDASQLPDFFGGTCTCADQGGCMRSDKGPWKDSEILK
                     MGRSGGTFCRHAGAFLSSDSQISSSDKPTYSLKVSDTSTAKSGSELEEMASPKTNTNN
                     HVPKLTPVSEYANGNISPTVLSEYEECVPMVDKVVDVAWQLQEMPNASEGPQYTSSLG
                     KIGSVRHIWSWLTAFFISFFTLLASLALPQTKEHSQLHSSSVRAELCDERIARESRPP
                     SPPRSTITERVIISSVLSRLGDLEKQIENLHSRKSEMPHEKEELLNAAVYRVDALEAE
                     LITTKKV"
     gene            70742. .73284
                     /gene="F14P1.3"
     CDS             join(70742. .70816,70957. .71382,72084. .72173,72904. .73284)
                     /gene="F14P1.3"
                     /note="Hypothetical protein"
                     /codon_start=1
                     /protein_id="AAF98406.1"
                     /db_xref="GI:9795588"
                     /translation="MEVMRVLHMNKGNGETSYAKNSTAQSNIISLGRRVMDEALKKLM
                     MSNSEISSIGIADLGCSSGPNSLLSISNIVDTIHNLCPDLDRPVPELRVSLNDLPSND
                     FNYICASLPEFYDRVNNNKEGLGFGRGGGESCFVSAVPGSFYGRLFPRRSLHFVHSSS
                     SLHWLSQKITGSHNRRELLSMGTPSSSSYVHGQRGIVGIIEEEKIDAFNAPYYAASSE
                     ELKMVIEKEGSFSIDRLEISPIDWEGGSISEESYDLVIRSKPEALASGRRVSNTIRAV
                     VEPMLEPTFGENVMDELFERYAKIVGEYFYVSSPRYAIVILSLVRAG"
     gene            74552. .76417
                     /gene="F14P1.4"
     CDS             join(74552. .74694,74769. .75090,75300. .75449,75558. .75735,
                     75908. .75987,76159. .76245,76328. .76417)
                     /gene="F14P1.4"
                     /note="Hypothetical protein"
                     /codon_start=1
                     /protein_id="AAF98400.1"
                     /db_xref="GI:9795582"
                     /translation="MFQDRKKSTAGVPPGSDGFPVIGETLQFMLSVNSGKGFYEFVRS
                     RRIRYGSCFRTSLFGETHVFLSTTESARLVLNNDSGMFTKRYIKSIGELVGDRSLLCA
                     PQHHHKILRSRLINLFSKRSTALMVRHFDELVVDALGGWEHRGTVVLLTDLLQITFKA
                     MCKMLVSLEKEEELGSMQRDVGFVCEAMLAFPLNLPWTRFHKGIMARGRVMEMLEKII
                     RERRNEINSHNNHHEDFLQQLLAVDNDTPQLTDAEIKDNILTMIIAGQDTTASALTWM
                     VKYLGENQKVLDILIEEQSQITKKASNKPFLELEDLSEMPYASKMVKESLRMASVVPW
                     FPRLVLQDCEMEGLL"
     gene            complement(77375. .78430)
                     /gene="F14P1.5"
     CDS             complement(77375. .78430)
                     /gene="F14P1.5"
                     /note="Unknown protein"
                     /codon_start=1
                     /protein_id="AAF98401.1"
                     /db_xref="GI:9795583"
                     /translation="MEAEKAEASWTYENLFQNCKYTVGLSAVVAFGYCCLPVIVNQLH
                     GFFIWFDWFTLVLRNLFFLSVVVSIIIFLIYVIFSKIQTTEKKEPDLYDQYVAAVPPP
                     STVTPFDSYNMSKFEDVGYYNNSTYRENVPAFQAVKAVEESSYSYDQEVVEMSSKSYR
                     RTSRSEKKMTERMVEYRKTETERVTKTGSWRSQSMDGLSSGEEFRRTIQTFVSEQKKM
                     HIIRASGVDDQLENGYIHRASGVDDRVHYWEHGAVPQWQNDDVPQLQNGGVPHWWQNG
                     SVSQLQNSDVPQWQNGGVLQLQNGDVPQWQTGGVSQLQNGVHEPQNGIVQWKGSRDDG
                     SGRHRHRHGHRSRKHRQ"
     gene            80606. .80776
                     /gene="F14P1.6"
     CDS             80606. .80776
                     /gene="F14P1.6"
                     /note="Similar to defensin; Similar to defensin AMP1 (anti
                     microbial protein)"
                     /codon_start=1
                     /protein_id="AAF98402.1"
                     /db_xref="GI:9795584"
                     /translation="MMAVEGRICERRSKTWTGFCGNTRGCDSQCKRWERASHGACHAQ
                     FPGFACFCYFNC"
     gene            complement(81552. .83365)
                     /gene="F14P1.7"
     CDS             complement(join(81552. .81800,81920. .82031,82136. .82233,
                     82329. .82436,82543. .82581,82670. .82741,82820. .83068,
                     83225. .83365))
                     /gene="F14P1.7"
                     /note="Unknown protein"
                     /codon_start=1
                     /protein_id="AAF98405.1"
                     /db_xref="GI:9795587"
                     /translation="MVAEALPKCPEAPLVLGLQPAALIDNVAPVDWSLLDQIPGDRGG
                     SIAVQKDELEHMLKELDAHISVAPLKKMAGGSVTNTVRGLSVGFGVATGIIGAYGDDE
                     QGQLFVSNMGFSGVSISRLRKKKGSTAQCVCLVDDSGNRTMRPCLSSAVKIQADELSK
                     EDFTGSKWLVLRYAVLNLQVIQAAIRFAKQEGLSVSLDLASFEMVRNSKSELRQLLES
                     GNIDLCFANEDEAAELLRGEQEAGPEAALEFLGRHCRWAVVTLGSKGCIAKHDKEVVH
                     ISAIGETVATDATGAGDLFASGFLYGLIKGLSLEECCKVGSCSGGSVIRALGGEVTPE
                     NWQWMHKQLQLKGLPVPDIHN"
     gene            complement(85358. .87513)
                     /gene="F14P1.8"
     CDS             complement(join(85358. .85648,86056. .86207,86310. .86435,
                     87001. .87144,87399. .87513))
                     /gene="F14P1.8"
                     /note="Unknown protein"
                     /codon_start=1

  Query Match             85.2%;  Score 1012;  DB 15;  Length 90341;
  Best Local Similarity   88.5%;  Pred. No. 8.6e-274;
  Matches 1175;  Conservative    0;  Mismatches    0;  Indels  153;  Gaps    1;

Qy          2 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57480 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 57539

Qy         62 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      57540 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 57599
```

```
Qy       122 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57600 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 57659

Qy       182 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57660 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 57719

Qy       242 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 301
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57720 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 57779

Qy       302 TTCTTGTAGCCCC---------------------------------------------- 314
            |||||||||||||
Db     57780 TTCTTGTAGCCCCACAGGTACATAATTTAACTATACAAAAAATAATTATGAATCATTTAA 57839

Qy       315 ------------------------------------------------------------ 314

Db     57840 AAGGATAACTTACGGTATGTAACATGTAACATGTAAAACAAAATGTACATTAATTGAGAT 57899

Qy       315 ----------------------------------------------ACAGTTGTGCAAAT 328
                                                          ||||||||||||||
Db     57900 TGATTGTTACGTGAAAAGATATAAGTTAATTTTATATTCTTGAATGACAGTTGTGCAAAT 57959

Qy       329 TATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCAT 388
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57960 TATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCAT 58019

Qy       389 CGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCAC 448
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58020 CGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCAC 58079

Qy       449 TCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAA 508
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58080 TCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAA 58139

Qy       509 CATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTA 568
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58140 CATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTA 58199

Qy       569 ACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGG 628
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58200 ACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGG 58259

Qy       629 ACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCAC 688
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58260 ACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCAC 58319

Qy       689 CATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAG 748
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58320 CATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAG 58379

Qy       749 CCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTT 808
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58380 CCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTT 58439

Qy       809 TTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGA 868
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58440 TTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGA 58499

Qy       869 GGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAG 928
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58500 GGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAG 58559

Qy       929 CGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCAC 988
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58560 CGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCAC 58619

Qy       989 CTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGA 1048
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58620 CTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGA 58679

Qy      1049 GGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTC 1108
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58680 GGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTC 58739

Qy      1109 ATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTA 1168
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58740 ATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTA 58799

Qy      1169 CCCCAAAA 1176
            ||||||||
Db     58800 CCCCAAAA 58807
```

RESULT 7
AC007797

```
LOCUS       AC007797               119942 bp    DNA     linear   PLN 12-SEP-2000
DEFINITION  Arabidopsis thaliana chromosome I BAC F6F9 genomic sequence,
            complete sequence.
ACCESSION   AC007797
VERSION     AC007797.7  GI:7839909
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 119942)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Miranda,M., Brooks,S.,
            Buehler,E., Chao,Q., Chin,C., Chiou,J., Choi,E., Gonzalez,A.,
            Howng,B., Johnson-Hopson,C., Khan,S., Kim,C., Koo,T., Lee,J.M.,
            Lenz,C., Liu,A., Liu,S., Mukharsky,N., Pham,P., Sakano,H.,
            Shinn,P., Toriumi,M., Vaysberg,M., Yu,G., Ecker,J., Theologis,A.
            and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 119942)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
            Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
            Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
            Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1999) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 119942)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Bei,Q., Buehler,E.,
            Chin,C., Chiou,J., Choi,E., Dunn,P., Gonzalez,A., Howng,B., Kim,C.,
            Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S.,
            Mukharsky,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P.,
            Thaveri,A., Toriumi,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J.,
            Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   4  (bases 1 to 119942)
  AUTHORS   Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
            Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.
            and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     On May 16, 2000 this sequence version replaced gi:7547094.
            Bases 60,267-119,942 of BAC clone F6F9 overlap with bases 1-59,676
            of IGF clone F14P1, AC024609 and bases 1-3,490 of BAC clone F6F9
            overlap with bases 89,221-92,710 of BAC clone T20H2, AC022472.
FEATURES             Location/Qualifiers
     source          1. .119942
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
```

```
                     /chromosome="I"
                     /clone="F6F9"
gene                 2448. .4393
                     /gene="F6F9.1"
CDS                  join(2448. .2640,2712. .2995,3068. .3502,3589. .3727,
                     3808. .3974,4064. .4393)
                     /gene="F6F9.1"
                     /note="Similar to endo-beta-1,4-glucanase"
                     /codon_start=1
                     /protein_id="AAG12562.1"
                     /db_xref="GI:10086502"
                     /translation="MVAKPRSRCCCCSVFIGVIILIAIIIAVIFTIRHRSNHSDDDGS
                     NVKNYANALKIAMQFFDIQKSGKLENNEISWRGDSGLKDGSEASIDLSKGLYDAGDHM
                     KFGFPMAFTATVLSWSILEYGDQMASLNLLDHAKDSLKWTTDFLINAHPSPNVLYIQV
                     GDPVTDHKCWDRPETMTRKRTLTKIDTKTPGTEVAAETAAAMAAASLVFKESDTKYSS
                     TLLKHAKQLFDFADNNRGSYSVNIPEVQSYYNSTGYGDELLWAASWLYHATEDQTYLD
                     FVSENGEEFGNFGSPSWFSWDNKLPGTHILLSRLTFFKKGLSGSKGLQGFKETAEAVM
                     CGLIPSSPTATSSRTDGGLIWVSEWNALQHPVSSAFLATLYSDYMLTSGVKELSCSDQ
                     SFKPSDLRKFARSQADYMLGKNPEKMSYLVGYGEKYPEFVHHRGASIPADATTGCKDG
                     FKWLNSDEPNPNVAYGALVGGPFLNDTFIDARNNSMQNEPSTYNSALVVGLLSSLVTT
                     SSSVESF"
gene                 6059. .7881
                     /gene="F6F9.2"
CDS                  join(6059. .6600,6686. .7091,7183. .7443,7547. .7633,
                     7747. .7881)
                     /gene="F6F9.2"
                     /codon_start=1
                     /product="sulfate adenylyltransferase"
                     /protein_id="AAG12541.1"
                     /db_xref="GI:10086481"
                     /translation="MSLMIRSSYVSHITLFQPRNSKPSSFTNQISFLSSSNNNPFLNL
                     VYKRNLTMQSVSKMTVKSSLIDPDGGELVELIVPETEIGVKKAESETMPKVKLNQIDL
                     EWVHVISEGWASPLKGFMREDEYLQSLHFNSLRLKNGTFVNMSLPIVLAIDDDTKEQI
                     GSSENVALVCPQGDIIGSLRSVEIYKHNKEERIARTWGTTSPGLPYVEEYITPSGNWL
                     IGGDLEVFEPIKYNDGLDHYRLSPKQLREEFDNRQADAVFAFQLRNPVHNGHALLMND
                     TRKRLLEMGYKNPVLLLHPLGGFTKADDVPLDVRMEQHSKVLEDGVLDPKTTIVSIFP
                     SPMHYAGPTEVQWHAKARINAGANFYIVGRDPAGMGHPTEKRDLYDPDHGKRVLSMAP
                     GLEKLNILPFRVAAYDTIEKKMAFFDPSRAKEFLFISGTKMRTYARTGENPPDGFMCP
                     SGWNVLVKYYESLQESEAKQQAVVSA"
gene                 complement(8394. .9399)
                     /gene="F6F9.3"
CDS                  complement(join(8394. .8526,8930. .9215,9321. .9399))
                     /gene="F6F9.3"
                     /codon_start=1
                     /product="vacuolar H+-pumping ATPase"
                     /protein_id="AAG12542.1"
                     /db_xref="GI:10086482"
                     /translation="MASTFSGDETAPFFGFLGAAAALVFSCMGAAYGTAKSGVGVASM
                     GVMRPELVMKSIVPVVMAGVLGIYGLIIAVIISTGINPKAKSYYLFDGYAHLSSGLAC
                     GLAGLSAGMAIGIVGDAGVRANAQQPKLFVGMILILIFAEALALYGLIVGIILSSRAG
                     QSRAE"
gene                 14164. .15678
                     /gene="F6F9.4"
CDS                  14164. .15678
                     /gene="F6F9.4"
                     /note="Unknown Protein"
                     /codon_start=1
                     /protein_id="AAG12543.1"
                     /db_xref="GI:10086483"
                     /translation="MHMQLLHNDRVVMYDRTNFGPSNISLPNGNCRDNPQDAVSKIDC
                     TAHSIEYDVATNTIRPLTVQSNTWCSSGSVRPDGVLVQTGGDRDGELKTRTFSPCNNN
                     QCDWVEMNNGLKKRRWYASNHILPDGKQIVMGGQGQFNYEFFPKTTNPNVVALPFLAE
                     THDQGQENNLYPFVFMNTDGNLFMFANNRAILLDYVKNTVVKTFPAIPGGDPRNYPST
                     GSAVLLPLKNLEADNVETEVLVCGGAPKGSYNLARKKTFVKALDTCARIKINDAKPEW
                     AVEKMPHARVMGDMIPLPNGDVLLINGGSFGTAAWELGRTPVLAPDLYHPENPVGSRF
                     ESLRPTTIPRMYHSAAILLRDGRVLVGGSNPHAFYNYTGVLFPTELSLEAFSPVYLQR
                     EFSNLRPKIISPEPQSMIKYGTNLKLKFSVTGEVTTPAKVTMVFPTFTTHSFAMNQRV
                     LVLDNVKFTRKGKSPMYEVQVRTPRSANIAWPGYYMIFVVNQDIPSEGVWVKLE"
gene                 complement(16924. .17682)
                     /gene="F6F9.5"
CDS                  complement(join(16924. .17094,17228. .17318,17417. .17496,
                     17611. .17682))
                     /gene="F6F9.5"
                     /codon_start=1
                     /product="Putative histone H3"
                     /protein_id="AAG12563.1"
                     /db_xref="GI:10086503"
                     /translation="MARTKQTARKSTGGKGPRKELATKAARKTRRPYRGGVKRAHRFR
                     PGTVALREIRKYQKSTDLLIRKLPFQRLVREIAQDFKVDLRFQSHAVLALQEAAEAYL
                     VGLFEDTNLCAIHAKRVTIMSKDIQLARRIRGERA"
gene                 18898. .22068
                     /gene="F6F9.7"
CDS                  join(18898. .19093,19182. .19243,19336. .19403,19510. .19631,
                     19701. .19824,19937. .19988,20276. .20393,20887. .21016,
                     21098. .21220,21301. .21354,21480. .21661,21734. .22068)
                     /gene="F6F9.7"
                     /note="Unknown Protein"
                     /codon_start=1
                     /protein_id="AAG12544.1"
                     /db_xref="GI:10086484"
                     /translation="MAEAMNSSEKKDESSEEKGGELLFCGATAWDIIGKRKGAMEGNL
                     VSPTRLRPLVGVNIRFVATGCASFHCVALDVEGRCYTWGRNEKGQLGHGDMIQRDRPT
                     VVSGLSKHKIVKAAAGRNHTVVVSDDGQSLGFGWNKYGQLGLGSAKNGFVSVEVESTP
                     LPCVVSDEVTNVACGADFTVWLSSTEGASILTAGLPQYGQLGHGTDNEFNMKDSSVRL
                     AYEAQPRPKAIASLAGETIVKVACGTNHTVFWSYIKTKFNLFGLSGGGQLYMWGKIKN
                     NGDDWMYPKPMMDLSGWNLRWMDSGSMHHFVGADSSCISWGHAQYGELGYGPNGQKSS
                     AAPKKVDMLEGMHVMGTDIADRLEQVTKAREPFFIRALNRKLKRYISSRFGISSSVSL
                     HCLLCLQLEVYDGKGSLEAESVEEVKEETLAPKQQAVTKRGASKKRKASKASSDSEQD
                     SDEDNSDKEKEVQGSDADSDYSEDGEEANGKKQSARGRGRGRGARGRGGRTSNGKAPP
                     VKTGGRRGRPRKS"
gene                 24177. .27256
                     /gene="F6F9.8"
CDS                  join(24177. .24269,24579. .25172,25331. .25510,25592. .25651,
                     25766. .27256)
                     /gene="F6F9.8"
                     /note="Unknown Protein"
                     /codon_start=1
                     /protein_id="AAG12559.1"
                     /db_xref="GI:10086499"
                     /translation="MGRSPASSCLRLISCSGGDDTSADPNSTALENKSSGDKRGWSFR
                     KKSGKQRGLITSVVSETTPASRTRETLESALLKSPSPDNNNVSEKQQQSFSVDEKKSQ
                     LPVVTYVAEPVDEKKTQSVIEEKTELLSVEEQIDHRTEVSPVIVESKGTETEEDDLIG
                     TELQGPSAADAAKIEEDVTSEVEMASKVEPEESESDDVIIVRKESDEKVDEKLDESVI
                     VVIQAAVRGFLARRELLRSKKVIKLQAAVRGHLVRSQAMGSLRCVQAIVKMQAMVRAR
                     HSTKDGSRVSATSDKSEPNAAAQKLLENKFAKHLMESTPKTKPINIKCDPTKPSSAWN
                     WLERWMSVPKPEKTSKANLTTEEQNLEETQNVKISPQADFVNSDSTVENKTETDMPSY
                     EASKVEGQNVELSETEKMSQYDSPEASAEVYYDSIQSQPLAAKPDSLLEEPEYVDGQI
                     KHSLKRKVSNPSFIAAQSKFEELTSSTGSNKAMTLSSKDDVLGEEGKTDIDSPDTTNT
                     IKDHSLEDVTPAELSGSECGTELSVTSSLDTLEKKSDAEGAEPRVEAKLLEDDTFKTD
                     QAELIEIDVKDATSMGTVEDPKEKVENAKDEVEISATHHEPVISTPDSKKRRAEDESG
                     PQAYALSEGALTPMTITESQATPASQASSSVKARKGKSEKSGSSQKRKVSKKITSSPK
                     QEIGTGEATEQEEGKEQKSGRRTSFGYDQEARESSGGKNSLPRFMQPTQSAKAKVQEH
                     NSPRSSPDLQERDVVSAKKRHSLPGVTNGKQVSPRIQRSASQAQQGTKGKNNLLQSFL
                     CNHSSQF"
gene                 28911. .31062
                     /gene="F6F9.9"
CDS                  join(28911. .28985,29480. .29677,29976. .30100,30198. .31062)
                     /gene="F6F9.9"
                     /note="Hypothetical Protein"
                     /codon_start=1
                     /protein_id="AAG12545.1"
                     /db_xref="GI:10086485"
                     /translation="MQENAPRTMSDVQRVLAISNKVFDMVRLFISEDSPSQVGSESQD
                     HLQAKSPLASHPSDDNLPPGFGGPFSANESQIKLSDIPLIKWKCSVQILLDREWKVVA
                     GDESKEVEAQNERELRVLEAFYPGASSIPPNPSVPADVEDSHHDDQQTIVIPILPVED
                     DDIAMDSASDFPTQSGVDVGTEPSITDENTSTSSTLPAGPDIMAALSAISNSKEQGSM
                     IDQDLLIKILSNPKLVENLVANRGSAGSVSSNTSSLYSSSTHEANGVVTTAPISSNGQ
                     FYAQPPITHIPPMAYTPHAPQDQPNYGAPPARDASYYKNLIQQHGGDRQETPPVQHLG
                     YRYNLQPGGGPNPEMVNSSNNNQRPRDSKPKIMKACMYFNSARGCRHGANCMYQHDAT
                     PYQPRNLNNGNINTSDMQNAKRMRFDRD"
gene                 complement(31753. .35330)
```

Query Match 85.2%; Score 1012; DB 15; Length 119942;
Best Local Similarity 88.5%; Pred. No. 8.6e-274;
Matches 1175; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

```
Qy          2 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117746 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 117805

Qy         62 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117806 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 117865

Qy        122 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117866 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 117925

Qy        182 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117926 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 117985

Qy        242 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117986 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 118045

Qy        302 TTCTTGTAGCCCC---------------------------------------------- 314
              |||||||||||||
Db     118046 TTCTTGTAGCCCCACAGGTACATAATTTAACTATACAAAAAATAATTATGAATCATTTAA 118105

Qy        315 ---------------------------------------------------------- 314

Db     118106 AAGGATAACTTACGGTATGTAACATGTAACATGTAAAACAAAATGTACATTAATTGAGAT 118165

Qy        315 ----------------------------------------------ACAGTTGTGCAAAT 328
                                                           |||||||||||||
Db     118166 TGATTGTTACGTGAAAAGATATAAGTTAATTTTATATTCTTGAATGACAGTTGTGCAAAT 118225

Qy        329 TATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCAT 388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118226 TATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCAT 118285

Qy        389 CGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCAC 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118286 CGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCAC 118345

Qy        449 TCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAA 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118346 TCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAA 118405

Qy        509 CATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTA 568
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118406 CATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTA 118465

Qy        569 ACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGG 628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118466 ACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGG 118525

Qy        629 ACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCAC 688
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118526 ACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCAC 118585

Qy        689 CATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAG 748
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118586 CATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAG 118645

Qy        749 CCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTT 808
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118646 CCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTT 118705

Qy        809 TTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGA 868
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118706 TTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGA 118765

Qy        869 GGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAG 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118766 GGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAG 118825

Qy        929 CGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCAC 988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118826 CGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCAC 118885

Qy        989 CTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGA 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118886 CTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGA 118945

Qy       1049 GGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTC 1108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118946 GGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTC 119005

Qy       1109 ATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTA 1168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119006 ATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTA 119065

Qy       1169 CCCCAAAA 1176
              ||||||||
Db     119066 CCCCAAAA 119073
```

RESULT 8
AX412268

```
LOCUS       AX412268                975 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 32 from Patent WO0222675.
ACCESSION   AX412268
VERSION     AX412268.1  GI:21444726
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
  TITLE     Plant genes, the expression of which are altered by pathogen
            infection
  JOURNAL   Patent: WO 0222675-A 32 21-MAR-2002;
            Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
            CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
            Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN
```

```
  Query Match             82.1%;  Score 975;  DB 6;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-263;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

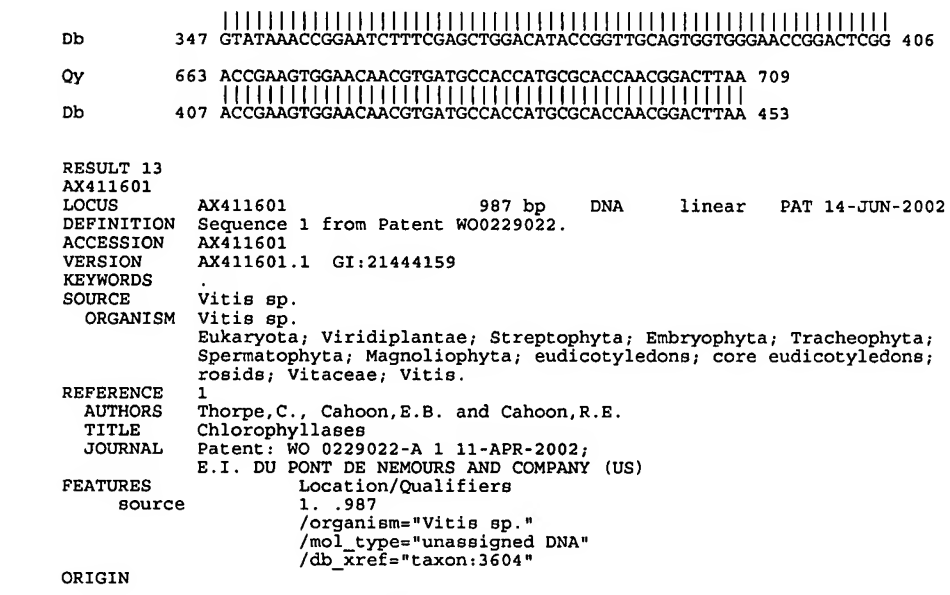
```
Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468

Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528

Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588

Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648

Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708

Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768

Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828

Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888

Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948

Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008

Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023

Db        961 GGTATCTTCGTCTAG 975
```

```
RESULT 9
AX412469
LOCUS       AX412469                975 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 233 from Patent WO0222675.
ACCESSION   AX412469
VERSION     AX412469.1  GI:21444927
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
  TITLE     Plant genes, the expression of which are altered by pathogen
            infection
  JOURNAL   Patent: WO 0222675-A 233 21-MAR-2002;
            Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
            CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
            Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN
```

```
 Query Match             82.1%;  Score 975;  DB 6;  Length 975;
 Best Local Similarity   100.0%;  Pred. No. 2.3e-263;
 Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108

Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168

Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228

Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288

Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348

Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408

Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468

Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528

Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588

Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648

Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708

Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768

Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828

Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888

Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948

Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008

Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023

Db        961 GGTATCTTCGTCTAG 975
```

```
RESULT 10
AX505463
LOCUS       AX505463                975 bp    DNA     linear   PAT 27-SEP-2002
DEFINITION  Sequence 158 from Patent WO0216655.
ACCESSION   AX505463
VERSION     AX505463.1  GI:23386700
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 158 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN

  Query Match             82.1%;  Score 975;  DB 6;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-263;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy         109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy         169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy         229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy         289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy         349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy         409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy         469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy         529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy         589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy         649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy         709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy         769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy         829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy         889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy         949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy        1009 GGTATCTTCGTCTAG 1023
               ||||||||||||||||
Db         961 GGTATCTTCGTCTAG 975


RESULT 11
AF337544
LOCUS       AF337544                1115 bp    mRNA    linear   PLN 22-OCT-2002
DEFINITION  Brassica oleracea chlorophyllase 1 mRNA, complete cds.
ACCESSION   AF337544
VERSION     AF337544.1  GI:24210532
KEYWORDS    .
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 1115)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Molecular characterization and differential regulation of three
            chlorophyllase genes, members of the novel serine esterase gene
            family, in broccoli
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1115)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2001) Institute of Botany, Academia Sinica,
            Taipei, Taiwan 11529, Republic of China
FEATURES             Location/Qualifiers
     source          1. .1115
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /cultivar="Green king"
                     /db_xref="taxon:3712"
                     /tissue_type="floret"
     5'UTR           1. .6
     CDS             7. .981
                     /EC_number="3.1.1.14"
                     /note="BoCLH1"
                     /codon_start=1
                     /product="chlorophyllase 1"
                     /protein_id="AAN51933.1"
                     /db_xref="GI:24210533"
                     /translation="MAGKEDSETFFSAATPLAFELGSLPTTVIPADPSATDLTAPPKP
                     VIITSPTVAGTYPVVLFFHGFYLRNYFYSDVINHVASHGYIVVAPQLCKILPPGGQVE
                     VDDAGKVINWTSKNLKAHLPSSVNANGNYTALVGHSRGGKTAFAVALGHAATLDPSIK
                     FSALVGIDPVAGISKCIRTDPEILTYKPESFDLDMPVAVIGTGLGPKSNMLMPPCAPA
                     EVNHEEFYIECKATKGHFVAADYGHMDMLDDNLPGFVGFMAGCMCKNGKRKKSEMRSF
                     VGGIVVAFLKYSIWGEMSEIRQILKDPSVSPARLDPSPELEEASGYLV"
     3'UTR           982. .1115
ORIGIN
```

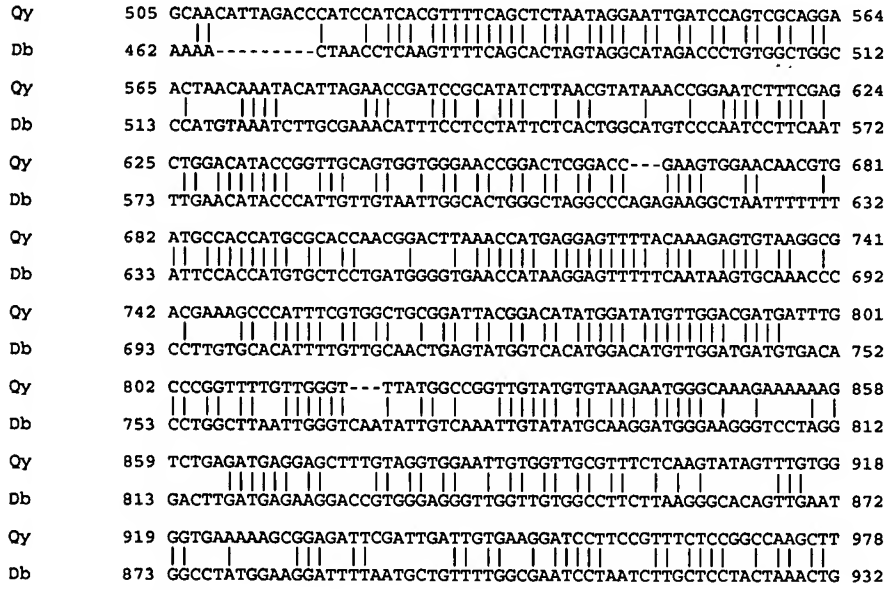
```
  Query Match             63.5%;  Score 754;  DB 15;  Length 1115;
  Best Local Similarity   82.7%;  Pred. No. 4.9e-201;
  Matches  862;  Conservative    0;  Mismatches  180;  Indels    0;  Gaps    0;

Qy         43 ATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCT 102
              |||||||||||| ||  |||||||||   |||||| ||| | || ||||||    |||
Db          1 ATACAAATGGCGGGGAAGGAGGACAGTGAGACGTTTTTCTCGGCGGCAACTCCTTTGGCG 60

Qy        103 TTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCA 162
              |||||| ||||||||| || |||||| ||| || |  || |||  || ||  ||||
Db         61 TTTGAGTTAGGCAGCCTTCCAACAACCGTGATCCCCGCAGACCCGTCGGCAACCGATTTG 120

Qy        163 ACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTC 222
              || ||||| ||||| || || | ||||||||  ||||| ||||||||||||| ||||||
Db        121 ACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAACCGTCGCCGGAACTTACCCCGTC 180

Qy        223 GTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCAC 282
              || |||||||||||||| || |||||||| ||||||||||||||| ||| ||||||||
Db        181 GTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTTCTACTCTGATGTTATTAACCAC 240

Qy        283 ATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGA 342
               | ||||| |||||| ||||||  ||||||||||||| ||| |||||  | |||||||||||
Db        241 GTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCTTTGCAAGATTTTGCCGCCGGGA 300

Qy        343 GGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAA 402
              |||||||||||||||||||||||||||   |||||||||||| | ||| ||||||||||
Db        301 GGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAACTGGACTTCGAAAAACCTCAAA 360

Qy        403 GCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGC 462
              ||||||| |||| ||| ||||| |||||||| || |||||| |||||||||| || |||
Db        361 GCTCACCTCCCAAGTTCAGTAAACGCTAATGGCAACTACACCGCACTCGTGGGTCATAGC 420

Qy        463 CGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCC 522
              ||||||| ||||| |||||||||||||||| ||||||| ||||||||| ||||||||||
Db        421 CGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCACGCCGCAACACTAGACCCATCC 480

Qy        523 ATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGA 582
              |||| |||||||||||| ||||||||   ||||||| ||||||  | |||||  ||| |||
Db        481 ATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGCAGGAATCAGCAAATGCATAAGA 540

Qy        583 ACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCA 642
              ||||||||  | ||||||||||| |||||||||| ||||| ||||||| |||||||||
Db        541 ACCGATCCCGAAATCTTAACGTACAAACCGGAATCATTCGACCTGGACATGCCGGTTGCA 600

Qy        643 GTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACG 702
              ||| | || || || |||||||||||  | ||||     |||||||||||||||||| ||
Db        601 GTGATCGGTACGGGTCTCGGACCGAAGAGTAACATGCTGATGCCACCATGCGCACCAGCG 660

Qy        703 GACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCT 762
              ||  | |||||||||||||||| |   |||||||||| ||||| |  |||||||||||
Db        661 GAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAGGCTACGAAGGGACATTTCGTGGCT 720

Qy        763 GCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATG 822
              ||||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||
Db        721 GCGGATTACGGACATATGGATATGTTGGACGATAATTTGCCCGGTTTTGTCGGGTTTATG 780

Qy        823 GCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGT 882
              || |||||||||||||||| ||  || | |||||| |||||||||  ||||||| |||
Db        781 GCGGGTTGTATGTGTAAGAACGGTAAACGCAAAAAGAGTGAGATGAGAAGCTTTGTTGGT 840

Qy        883 GGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTG 942
              ||||||||||||||||||| ||||||||| | ||||||||||    | ||||||||  |
Db        841 GGAATTGTGGTTGCGTTTCTAAAGTATAGTATATGGGGTGAAATGTCAGAGATTCGACAG 900

Qy        943 ATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAA 1002
              ||| ||||||||||||| |||||||| || | |||||||||||| || ||| ||||||
Db        901 ATTTGAAGGATCCTTCTGTTTCTCCAGCGAGGCTTGATCCTTCGCCGGAGCTGGAAGAG 960

Qy       1003 GCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATAT 1062
              ||||||||    ||||||||||||||| |         ||| |          |  ||  |
Db        961 GCTTCTGGTTATCTCGTCTAGATTTGTGGTTATGTAACGTATTGGTATCCCCAGAAAGAC 1020

Qy       1063 TTGAAAAACCTATCAATGTTTT 1084
              |  |||  | ||  ||  || |
Db       1021 TCCAAACTCGTAATAAACTTAT 1042

RESULT 12
BX510604
LOCUS       BX510604                 453 bp    DNA     linear   STS 10-JUN-2003
DEFINITION  Arabidopsis thaliana transposon insertion STS SM_3.40564, sequence
            tagged site.
ACCESSION   BX510604
VERSION     BX510604.1  GI:30577550
KEYWORDS    STS; STS, sequence tagged site.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
            Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 453)
  AUTHORS   Clarke,J.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2003) Clarke J.H., John Innes Centre, Colney
            Lane, Norwich, NR4 7UJ, UK
COMMENT     AT denotes an activation tag dissociation transposon within a
            single line, ET an enhancer trap dissociation transposon, GT a gene
            trap dissociation transposon, MT a mis-expression enhancer trap
            dissociation transposon, SM a defective suppressor mutator
            transposon. _3 denotes a sequence derived from the 3'end of the
            transposon, _5 denotes a sequence derived from the 5'end of the
            transposon BBSRC GARNet, ATIS project
            On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
            code: N127275.
FEATURES             Location/Qualifiers
     source          1..453
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /variety="Columbia-0 NASC stock code N1092"
                     /db_xref="taxon:3702"
                     /clone="AC024609"
                     /note="Derived from superpool 19.36 NASC code N40883"
     STS             1..453
                     /standard_name="SM_3.40564"
ORIGIN

  Query Match             33.3%;  Score 395.8;  DB 10;  Length 453;
  Best Local Similarity   98.3%;  Pred. No. 5.2e-100;
  Matches  400;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy        303 TCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGA 362
               | || | |     |||||||||||||||||||||||||||||||||||||||||||||
Db         47 TATTCTTGAATGACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGA 106

Qy        363 TGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGT 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 TGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGT 166

Qy        423 AAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTT 226

Qy        483 TGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAAT 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 TGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAAT 286

Qy        543 AGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAAC 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAAC 346

Qy        603 GTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGG 662
```



```
Db        347 GTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGG 406

Qy        663 ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAA 709
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        407 ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAA 453


RESULT 13
AX411601
LOCUS       AX411601                 987 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 1 from Patent WO0229022.
ACCESSION   AX411601
VERSION     AX411601.1  GI:21444159
KEYWORDS    .
SOURCE      Vitis sp.
  ORGANISM  Vitis sp.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; Vitaceae; Vitis.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 1 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .987
                     /organism="Vitis sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3604"
ORIGIN

  Query Match             20.2%;  Score 239.4;  DB 6;  Length 987;
  Best Local Similarity   57.6%;  Pred. No. 6.7e-56;
  Matches  473;  Conservative    0;  Mismatches  336;  Indels   12;  Gaps    2;

Qy        167 CACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTT 226
              | || || || ||  ||   ||       |||||  |     || || || || ||  | |
Db         26 CCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTCTCT 85

Qy        227 TATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCG 286
              | ||  | ||||||||  | || ||||||  |||||| || |  | ||| | | |||
Db         86 TGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCATTT 145

Qy        287 CTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGC 346
              |||| ||||| | |||| | || || || ||||| | |  | |||| || ||   ||
Db        146 CTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTGGAA 205

Qy        347 AAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTC 406
              ||  |  |    || | |   || | ||| |||  |||   |  |||  || ||
Db        206 TTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAATCTG 265

Qy        407 ACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCG 466
              || |||      || ||  |  |    |      | ||    |||||||| | |
Db        266 TGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCAGAG 325

Qy        467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCA 526
              | || || || || |||||  | || |||||  ||||          ||  ||||| |||
Db        326 GGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGC---------TGATACATCCCTCA 376

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
              || ||||| ||| ||||| | || || || |  ||      | ||||       ||| |
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        587 ATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGG 646
              |||  | ||| |||| |||   ||  | || ||| | || |  || || |||   ||
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAGCAATCCCAGTTTGCGTAA 496

Qy        647 TGGGAACCGGACTCGGACCGAAG---TGGAACAACGTGATGCCACCATGCGCACCAACGG 703
              | || || ||  | ||     ||    |||||  | | |     |||| || |||   |
Db        497 TCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAGATG 556

Qy        704 ACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTG 763
```

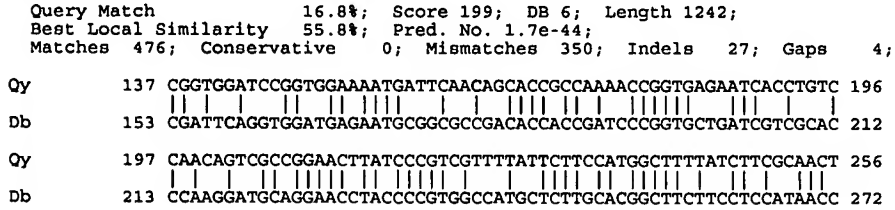


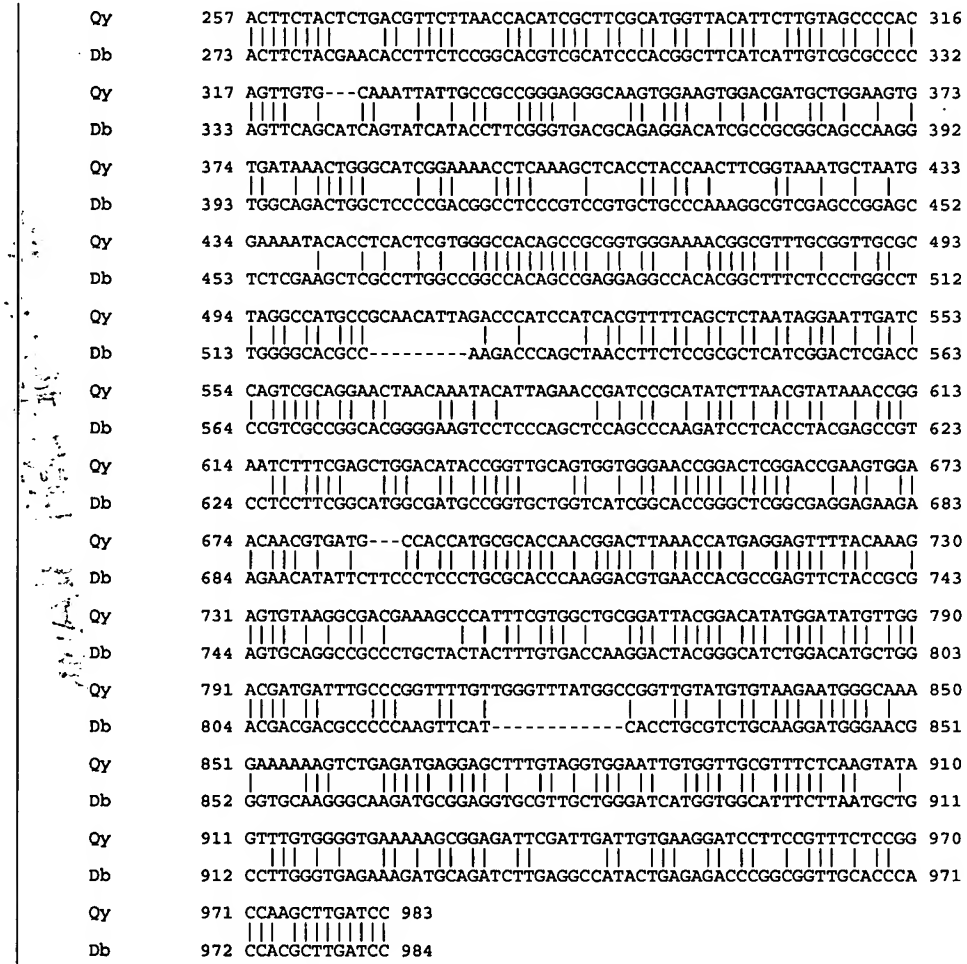
```
              | |||||||  ||||||| ||  ||||||||  |  |     |||| || ||| ||
Db        557 GAGTGAACCATGTAGAGTTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGACTA 616

Qy        764 CGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGG 823
              | || || ||  |||| ||||| || ||| || |  | || |   |||||    ||
Db        617 CTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGATTT 676

Qy        824 CCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTG 883
              | |||| ||| || |||| |||| |  |     | |       |||||||| | ||| ||||
Db        677 CGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGGGTG 736

Qy        884 GAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGA 943
              |  || | ||||| ||  | |||  |  ||||   ||| | |  |  ||  |   |  |
Db        737 GCCTTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATTTCAAAGCCA 796

Qy        944 TTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCT 984
              ||||  | || |||    |  |||| |  ||||||||||||
Db        797 TTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCT 837


RESULT 14
AX411607
LOCUS       AX411607                1174 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 7 from Patent WO0229022.
ACCESSION   AX411607
VERSION     AX411607.1  GI:21444162
KEYWORDS    .
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 7 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1174
                     /organism="Glycine max"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3847"
ORIGIN

  Query Match             17.1%;  Score 202.6;  DB 6;  Length 1174;
  Best Local Similarity   56.9%;  Pred. No. 1.6e-45;
  Matches  480;  Conservative    0;  Mismatches  339;  Indels   24;  Gaps    5;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
              |||| |||  |  |||| ||||||||  ||  ||||      |||| ||  | ||  |
Db        102 AATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTGGCTCA 161

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
              || || ||  | || |||| |||||| ||||  ||||||||   || |||||||||  |
Db        162 TACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTGAGCTC 221

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTG 333
              ||   |||||| ||||| ||||| | |||  |||| || || ||| ||||       |
Db        222 CTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGAGTGTA 281

Qy        334 CCGCC------GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGG--- 384
              | | |      || |      |  |||||  |   ||| || |  ||  |  | |||
Db        282 CGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATTGGCTA 341

Qy        385 GCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACC 444
              ||   |||    ||  || |||  || |||       ||  | || ||    |  |
Db        342 GCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATAAATTG 401

Qy        445 TCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCC 504
              |     || ||||||  |||| |||||  |  |||| || || ||  ||||
Db        402 GTTTTATCAGGTCACAGCAAGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTTATGCT 461
```



```
RESULT 15
AX411617
LOCUS       AX411617               1242 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 17 from Patent WO0229022.
ACCESSION   AX411617
VERSION     AX411617.1  GI:21444167
KEYWORDS    .
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 17 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1242
                     /organism="Triticum aestivum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4565"
ORIGIN
```

Query Match 16.8%; Score 199; DB 6; Length 1242;
Best Local Similarity 55.8%; Pred. No. 1.7e-44;
Matches 476; Conservative 0; Mismatches 350; Indels 27; Gaps 4;





Search completed: March 20, 2006, 16:11:51
Job time : 4208.73 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:18 ; Search time 5384.1 Seconds
(without alignments)
10323.556 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		849.8	71.5	1169	4	CNS0ADGI	BX814036 Arabidops
2		772.8	65.1	796	5	BU636473	BU636473 006A06 In
3		697.4	58.7	699	6	CB255908	CB255908 12-E01274
c 4		605.6	51.0	715	2	BE039090	BE039090 AB09E05 A
c 5		548	46.1	563	6	CF773835	CF773835 AG_FSL_23
6		533.8	44.9	681	10	CW836771	CW836771 ET8513.Ds
7		525	44.2	525	1	AV525625	AV525625 AV525625
8		524	44.1	741	8	DN777869	DN777869 E5668 [C4
c 9		506.6	42.6	794	9	B20290	B20290 T21K8-T7 TA
c 10		493.8	41.6	547	1	AI997409	AI997409 701553812
11		413	34.8	413	1	AV561535	AV561535 AV561535
c 12		412.4	34.7	570	8	DN779181	DN779181 E7566 [C4
c 13		403	33.9	547	8	DN779131	DN779131 E7515 [C4
c 14		391.6	33.0	430	3	BP605375	BP605375 BP605375
c 15		389.4	32.8	411	1	AV817693	AV817693 AV817693
16		387.2	32.6	499	8	DN778168	DN778168 E6120 [C4
c 17		385	32.4	396	1	AV535509	AV535509 AV535509
c 18		372	31.3	404	1	AV440587	AV440587 AV440587
c 19		366.8	30.9	720	9	BZ427710	BZ427710 BONOW28TR
c 20		356.2	30.0	430	3	BP648469	BP648469 BP648469
c 21		349.4	29.4	403	3	BP660185	BP660185 BP660185
22		348.4	29.3	392	3	BP830895	BP830895 BP830895
23		344.4	29.0	495	8	CX189147	CX189147 79-E01112
c 24		339	28.5	339	1	AV563620	AV563620 AV563620
25		328	27.6	402	3	BP823922	BP823922 BP823922
26		327.6	27.6	378	3	BP835886	BP835886 BP835886
27		323.2	27.2	474	8	DN777868	DN777868 E5667 [C4
c 28		310	26.1	371	10	CW836770	CW836770 ET8513.Ds
c 29		309.6	26.1	431	3	BP625285	BP625285 BP625285
c 30		287.6	24.2	392	3	BP653864	BP653864 BP653864
c 31		286	24.1	973	10	CL478966	CL478966 SAIL_2_F0
32		285.2	24.0	387	3	BP833118	BP833118 BP833118
33		283.6	23.9	387	3	BP821899	BP821899 BP821899
c 34		279.6	23.5	825	10	CL478965	CL478965 SAIL_2_F0
35		271.8	22.9	366	3	BP857816	BP857816 BP857816
c 36		249.2	21.0	373	3	BP655953	BP655953 BP655953
37		239.4	20.2	316	9	BZ292665	BZ292665 SALK_1249
38		219	18.4	959	10	CL468490	CL468490 SAIL_1287
39		217.2	18.3	668	7	CV187182	CV187182 Mdlv2_401
c 40		204.8	17.2	764	9	BZ070295	BZ070295 lkj38h09.
41		202.6	17.1	713	8	CX022963	CX022963 Mdlv4-404
42		202.4	17.0	799	8	CX046588	CX046588 UCRCS07_9
43		201	16.9	837	8	CX046450	CX046450 UCRCS07_8
44		195	16.4	401	9	BZ290866	BZ290866 SALK_0929
45		187	15.7	879	8	CX046657	CX046657 UCRCS07_9

ALIGNMENTS

RESULT 1
CNS0ADGI
LOCUS CNS0ADGI 1169 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTPB53ZC07 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).
ACCESSION BX814036
VERSION BX814036.1 GI:42474034
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1169)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1169)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length librairies construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES Location/Qualifiers
source 1. .1169
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

```
                     /db_xref="taxon:3702"
                     /clone="GSLTFB53ZC07"
                     /tissue_type="Flowers and buds"
                     /ecotype="Col-0"
                     /plasmid="pCMVSPORT_6"
     gene            complement(1. .1169)
                     /gene="At1g19670"
ORIGIN

  Query Match             71.5%;  Score 849.8;  DB 4;  Length 1169;
  Best Local Similarity   86.7%;  Pred. No. 4.8e-238;
  Matches 1014;  Conservative    0;  Mismatches    2;  Indels  153;  Gaps    1;

Qy         25 CTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCT 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCT 60

Qy         85 GTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGAT 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGAT 120

Qy        145 CCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTC 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTC 180

Qy        205 GCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTAC 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTAC 240

Qy        265 TCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCC---------- 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGGTACAT 300

Qy        315 ------------------------------------------------------------ 314

Db        301 AATTTAACTATACAAAAAATAATTATGAATCATTTAAAAGGATAACTTACGGTATGTAAC 360

Qy        315 ------------------------------------------------------------ 314

Db        361 ATGTAACATGTAAAACAAAATGTACATTAATTGAGATTGATTGTTACGTGAAAAGATATA 420

Qy        315 ----------------------ACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTG 351
                                     ||||||||||||||||||||||||||||||||||||
Db        421 AGTTAATTTTATATTCTTGAATGACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTG 480

Qy        352 GAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTA 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTA 540

Qy        412 CCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGG 471
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        541 CCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCTACAGCCGCGGTGGG 600

Qy        472 AAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTT 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTT 660

Qy        532 TCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCG 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCG 720

Qy        592 CATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGA 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGA 780

Qy        652 ACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAAC 711
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCGTACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAAC 840

Qy        712 CATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTAC 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTAC 900

Qy        772 GGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGT 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGT 960

Qy        832 ATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTG 891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTG 1020

Qy        892 GTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAG 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAG 1080

Qy        952 GATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGT 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGT 1140

Qy       1012 ATCTTCGTCTAGATTTGTGTTATGTACTA 1040
              ||||||||||||||||||||||||||||||
Db       1141 ATCTTCGTCTAGATTTGTGTTATGTACTA 1169


RESULT 2
BU636473
LOCUS       BU636473                796 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  006A06 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
            sequence.
ACCESSION   BU636473
VERSION     BU636473.1  GI:23303728
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
            and Welinder,K.G.
  TITLE     EST sequencing of Erysiphe cichoracearum infected Arabidopsis
            plants
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Karen G. Welinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /dev_stage="Plant 3 weeks old, three days post infection"
                     /clone_lib="Infected Arabidopsis Leaf"
                     /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA
                     library of Arabidopsis and E. cichoracearum infected leaf
                     from three weeks old Arabidopsis plants. Plants were
                     harvested 3 days after infection and mRNA oligo dT
                     selected."
ORIGIN

  Query Match             65.1%;  Score 772.8;  DB 5;  Length 796;
  Best Local Similarity   99.5%;  Pred. No. 1.9e-215;
  Matches  786;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    1;

Qy         30 ATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGT 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 ATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGT 66

Qy         90 AACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGT 149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGT 126
```

```
Qy        150 GGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGG 209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGG 186

Qy        210 AACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGA 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGA 246

Qy        270 CGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTG--TAGCCCCACAGTTGTGCAAA 327
              |||||||||||||||||||||||||||||||||||||  ||||||  ||||||||||||||||||||
Db        247 CGTTCTTAACCACATCGCTTCGCATGGTTACCTTCTTGGTTAGCCCCACAGTTGTGCAAA 306

Qy        328 TTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCA 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCA 366

Qy        388 TCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCA 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 TCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCA 426

Qy        448 CTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCA 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCA 486

Qy        508 ACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACT 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACT 546

Qy        568 AACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 AACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTG 606

Qy        628 GACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 GACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCA 666

Qy        688 CCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAA 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAA 726

Qy        748 GCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGT 807
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 GCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGT 786

Qy        808 TTTGTTGGGT 817
              |||||  ||||
Db        787 TTTGTAGGGT 796
```

```
RESULT 3
CB255908
LOCUS       CB255908                699 bp    mRNA    linear   EST 06-NOV-2003
DEFINITION  12-E012744-027-007-H04-T7R MPIZ-ADIS-027 Arabidopsis thaliana cDNA
            clone MPIZp772H047Q 5-PRIME, mRNA sequence.
ACCESSION   CB255908
VERSION     CB255908.1  GI:32880681
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
            Mitchell-Olds,T. and Weisshaar,B.
  TITLE     Large-scale identification and analysis of genome-wide
            single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
  JOURNAL   Genome Res. 13 (6), 1250-1257 (2003)
   PUBMED   12799357
COMMENT     Contact: Weisshaar B
            ADIS DNA core facility at MPIZ
            Max-Planck-Institute for Plant Breeding Research
            Carl-von-Linne Weg 10, 50829 Koeln, Germany
            Fax: 00492215062851
            Email: weisshaa@mpiz-koeln.mpg.de
            Insert Length: 699   Std Error: 0.00
            Plate: 7  row: H  column: 04
            Seq primer: T7R; CTAATACGACTCACTATAGGGA.
FEATURES             Location/Qualifiers
     source          1. .699
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Ei-2"
                     /db_xref="GABI:592398"
                     /db_xref="taxon:3702"
                     /clone="MPIZp772H047Q"
                     /tissue_type="whole plant"
                     /dev_stage="adult plant, mixed stresses"
                     /lab_host="E. coli TOP10"
                     /clone_lib="MPIZ-ADIS-027"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; cDNA
                     library from Arabidopsis thaliana, accession Eifel-2; ten
                     week old total plants grown under long-day conditions in
                     soil, whole adult plants were treated for 24 hours with
                     different stresses, (1) at 4M-0 C in the dark, (2), at 37
                     Grad C in the dark, (3) lying in the lab after removing
                     from soil, (4) in the greenhouse after wounding leaves
                     with a foreceps, (5) in the lab watering with a 150 mM
                     NaCl solution, (6) at 26 M-0C in the light/UV; equal
                     quantities of stressed plant material were pooled; library
                     was made at the Max-Planck-Institute for Plant Breeding
                     Research, Cologne, Germany; cloning sites SalI-NotI,
                     primer sites and orientation:
                     T7-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
                     Sequencing granted in the context of the GABI Arabidopsis
                     Verbund I: Genetic Diversity, 'Establishment of
                     high-efficiency SNP-based mapping tools and development of
                     methods for genome-wide mutation detection' PI: Bernd
                     Weisshaar Sequence submission managed by RZPD/GABI-Primary
                     database: http://gabi.rzpd.de. This clone is available
                     from RZPD; contact RZPD (clone@rzpd.de) for further
                     information."
ORIGIN
```

Query Match 58.7%; Score 697.4; DB 6; Length 699;
Best Local Similarity 99.9%; Pred. No. 2.7e-193;
Matches 698; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         18 CACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTT 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTT 60

Qy         78 TTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACC 120

Qy        138 GGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCC 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCC 180

Qy        198 AACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTA 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTA 240

Qy        258 CTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACA 300

Qy        318 GTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGAT 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGAT 360

Qy        378 AAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAA 420
```

```
Qy        438 ATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGG 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGG 480

Qy        498 CCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGT 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGT 540

Qy        558 CGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATC 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATC 600

Qy        618 TTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAA 677
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        601 TTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCCGACTCGGACCGAAGTGGAACAA 660

Qy        678 CGTGATGCCACCATGCGCACCAACGGACTTAAACCATGA 716
              ||||||||||||||||||||||||||||||||||||||||
Db        661 CGTGATGCCACCATGCGCACCAACGGACTTAAACCATGA 699


RESULT 4
BE039090/c
LOCUS       BE039090                715 bp    mRNA    linear   EST 07-JUN-2000
DEFINITION  AB09E05 AB Arabidopsis thaliana cDNA 5' similar to
            coronatine-induced protein 1, mRNA sequence.
ACCESSION   BE039090
VERSION     BE039090.1  GI:8334106
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 715)
  AUTHORS   Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M.,
            Ferrea,H., Kawasaki,S., McCollough,A., Michalowski,C.B.,
            Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
  TITLE     Functional Genomics of Plant Stress Tolerance
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Michalowski,C.B.
            University of Arizona
            Bio Sciences West room 513, Tucson, AZ 85721, USA
            Tel: 520-621-7982
            Fax: 520-621-1697
            Email: cbm@u.arizona.edu.
FEATURES             Location/Qualifiers
     source          1. .715
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /tissue_type="leaves, seedlings"
                     /dev_stage="2-3 weeks"
                     /clone_lib="AB"
                     /note="200mM NaCl"
ORIGIN

  Query Match             51.0%;  Score 605.6;  DB 2;  Length 715;
  Best Local Similarity   97.4%;  Pred. No. 2.7e-166;
  Matches  647;  Conservative    0;  Mismatches   14;  Indels    3;  Gaps    3;

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
              |||||||||| ||||| ||||| |||||||||||||||| ||||| |||||||| |||
Db        675 CGTTTTCAGCTTTAATA-GAATTTATCCAGTCGCAGGAACTACCAAATCCATTAGAGCCG 617

Qy        587 ATCC-GCATATCTTAACGTATAAACCGGA-ATCTTTCGAGCTGGACATACCGGTTGCAGT 644
              ||| |||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db        616 TTCCGGCATATCTTAACGTATAAACCGGAGATCTTTCGAGCTGGACAAACCGGTTGCAGT 557

Qy        645 GGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGA 704
              ||||| |||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        556 GGTGGAAACCGGACTCGGACCGAAGTGGAACAACGTAATGCCACCATGCGCACCAACGGA 497

Qy        705 CTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGC 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGC 437

Qy        765 GGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGC 824
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        436 GGATTACGGACATATGGATATGTTGGACAATGATTTGCCCGGTTTTGTTGGGTTTATGGC 377

Qy        825 CGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGG 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 CGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGG 317

Qy        885 AATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGAT 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 AATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGAT 257

Qy        945 TGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGC 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 TGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGC 197

Qy       1005 TTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTT 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 TTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTT 137

Qy       1065 GAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCAT 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCAT 77

Qy       1125 GTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAA 1184
              |||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db         76 GTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTAAAAAAAAAAAAAAAA 17

Qy       1185 AAAA 1188
              ||||
Db         16 AAAA 13


RESULT 5
CF773835/c
LOCUS       CF773835                563 bp    mRNA    linear   EST 14-APR-2004
DEFINITION  AG_FSL_23D02 Arabidopsis ag-1 35S:AG-GR forward subtraction library
            Arabidopsis thaliana cDNA clone 23D02, mRNA sequence.
ACCESSION   CF773835
VERSION     CF773835.1  GI:46368801
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 563)
  AUTHORS   Hu,W., Wang,Y., Bowers,C. and Ma,H.
  TITLE     Isolation, sequence analysis, and expression studies of florally
            expressed cDNAs in Arabidopsis
  JOURNAL   Plant Mol. Biol. 53 (4), 545-563 (2003)
COMMENT     Contact: Hong Ma
            Hong Ma lab
            The Pennsylvania State University
            315 Wartik Laboratory, University Park, PA 16802, USA
            Tel: 8148636414
            Fax: 8148631357
            Email: hxm16@psu.edu
            Seq primer: M13 Universal
            High quality sequence stop: 563
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .563
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Landsberg erecta"
                     /db_xref="taxon:3702"
```

/clone="23D02"
/tissue_type="Inflorescence lacking open or older flowers"
/dev_stage="4 week old plants"
/clone_lib="Arabidopsis ag-1 35S:AG-GR forward subtraction library"
/note="Vector: The pT-Adv vector in the AdvanTAge Cloning Kit(CLONETECH); Subtractive library constructed using the CLONTECH PCR-select cDNA subtraction kit. Transgenic plants homozygous for the ag-1 mutation and carrying a 35S:AG-GR construct were generated. RNA was isolated from inflorescences with floral meristems and unopened floral buds and without open or older flowers from transgenic plants that were treated with dexamethasone for 3 days, and used for synthesis of the tester cDNA. RNA from the inflorescences of transgenic plants treated with a control solution without dexamethasone was used to synthesize the driver cDNA. It was observed that a 3-day treatment with dexamethasone can induce anther and carpel characteristics in the flowers of the ag-1 35S:AG-GR plants."

ORIGIN

Query Match 46.1%; Score 548; DB 6; Length 563;
Best Local Similarity 99.1%; Pred. No. 2.2e-149;
Matches 551; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 TTGCGGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 497

Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        496 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGACCCGATCCGCATATCTTA 437

Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 ATGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 377

Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 317

Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 257

Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 197

Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 137

Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 77

Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 17

Qy       1021 TAGATTTGTGTTATGT 1036
              ||||||||||||||||
Db         16 TAGATTTGTGTTATGT 1
```

RESULT 6
CW836771
LOCUS CW836771 681 bp DNA linear GSS 26-NOV-2004
DEFINITION ET8513.Ds5.05.20.01.JU94.b.681 Arabidopsis thaliana Landsberg Ds insertion lines Arabidopsis thaliana genomic clone ET8513, genomic survey sequence.
ACCESSION CW836771
VERSION CW836771.1 GI:56092561
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 681)
AUTHORS May,B.P., Simorowski,J., Arroyo,J.-M., Vaughn,M.W., Shen,R., McCombie,W.R. and Martienssen,R.A.
TITLE Arabidopsis genomic sequences flanking Ds enhancer and gene traps in transgenic lines
JOURNAL Unpublished (2004)
COMMENT Contact: Martienssen RA
Cold Spring Harbor Laboratory
1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8322
Fax: 516 367 8369
Email: martiens@cshl.org
This sequence flanks a Ds transposon carrying an enhancer trap in line ET8513. The transposon is located within At1g19670.
Class: transposon-tagged.
FEATURES Location/Qualifiers
source 1. .681
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Landsberg erecta"
/db_xref="taxon:3702"
/clone="ET8513"
/clone_lib="Arabidopsis thaliana Landsberg Ds insertion lines"
/note="Lines of Arabidopsis thaliana were generated which each contain a Ds transposon carrying a glucuronidase reporter gene. Genomic DNA flanking the transposon insertion in each line was amplified by PCR and directly sequenced. More information is available at http://genetrap.cshl.edu."

ORIGIN

Query Match 44.9%; Score 533.8; DB 10; Length 681;
Best Local Similarity 98.7%; Pred. No. 3.5e-145;
Matches 538; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy        303 TCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGA 362
              | || | |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 TATTCTTGAATGACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGA 196

Qy        363 TGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGT 422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGT 256

Qy        423 AAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTT 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 AAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTT 316

Qy        483 TGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAAT 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 TGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAAT 376

Qy        543 AGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAAC 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        377 AGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAAC 436

Qy        603 GTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGG 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        437 GTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGG 496

Qy        663 ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTT 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTT 556

Qy        723 TTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGA 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        557 TTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGA 616

Qy        783 TATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAA 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 TATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAA 676

Qy        843 TGGGC 847
              |||||
Db        677 TGGGC 681


RESULT 7
AV525625
LOCUS       AV525625                525 bp    mRNA    linear   EST 18-FEB-2004
DEFINITION  AV525625 Arabidopsis thaliana aboveground organs two to six-week
            old Arabidopsis thaliana cDNA clone APD27d04R 5', mRNA sequence.
ACCESSION   AV525625
VERSION     AV525625.1  GI:8685153
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 525)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7 (3), 175-180 (2000)
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .525
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="APD27d04R"
                     /tissue_type="aboveground organs"
                     /dev_stage="two to six-week old"
                     /clone_lib="Arabidopsis thaliana aboveground organs two to
                     six-week old"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match             44.2%;  Score 525;  DB 1;  Length 525;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-142;
  Matches  525;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 CAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTT 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTT 60

Qy         80 CCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGG 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGG 120

Qy        140 TGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAA 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAA 180

Qy        200 CAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACT 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACT 240

Qy        260 TCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGT 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 TCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGT 300

Qy        320 TGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAA 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAA 360

Qy        380 ACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAAT 439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAAT 420

Qy        440 ACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCC 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCC 480

Qy        500 ATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAG 544
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAG 525


RESULT 8
DN777869
LOCUS       DN777869                741 bp    mRNA    linear   EST 05-APR-2005
DEFINITION  E5668 [C4091|slt12_e11] Salinity Library Thellungiella salsuginea
            cDNA, mRNA sequence.
ACCESSION   DN777869
VERSION     DN777869.1  GI:62206750
KEYWORDS    EST.
SOURCE      Thellungiella salsuginea
  ORGANISM  Thellungiella salsuginea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   Wong,C.E., Li,Y., Whitty,B.R., Diaz-Camino,C., Akhter,S.R.,
            Brandle,J.E., Golding,G.B., Weretilnyk,E.A., Moffatt,B.A. and
            Griffith,M.
  TITLE     Expressed sequence tags from the Yukon ecotype of Thellungiella
            reveal that gene expression in response to cold, drought and
            salinity shows little overlap
  JOURNAL   Plant Mol. Biol. 58 (4), 561-574 (2005)
   PUBMED   16021339
COMMENT     Contact: Moffatt,B
            Department of Biology
            University of Waterloo
            200, University Ave West,Ontario, Canada N2L 3G1
            Tel: (519) 888-4567 ext 2517
            Fax: (519) 746-0614
            Email: moffatt@uwaterloo.ca.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Thellungiella salsuginea"
                     /mol_type="mRNA"
                     /cultivar="Yukon"
                     /db_xref="taxon:72664"
                     /tissue_type="Leave"
                     /dev_stage="4-week old plants"
                     /clone_lib="Salinity Library"
                     /note="Vector: pCMVSPORT6; Site_1: Not1; Site_2: Sal1;
                     Library was prepared following the protocol for Invitrogen
                     SuperScript Plasmid System  with GAteway Technology for
                     cDNA synthesis and cloning"
ORIGIN

  Query Match             44.1%;  Score 524;  DB 8;  Length 741;
  Best Local Similarity   85.5%;  Pred. No. 2.7e-142;
  Matches  595;  Conservative    0;  Mismatches  100;  Indels    1;  Gaps    1;

Qy         39 TTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGC 98
              | | | | ||||||||||||| || |||| |   ||||||| ||| | || ||||||||
Db         14 TATCAAAAAAAATGGCGGCGATGGATGACAATGAGACGTTTTTCTCGGCGGCAACTCCGGT 73

Qy         99 GGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGA 158
              ||| ||||||| |||||||| ||||| |||| |||  ||| |||||   ||||| ||
```

```
Db        74 GGCGTTTGAGAGAGGCAGCCTTCCGACCACCGTGATCACGGCGGATCCTTCGGAAACCGA 133

Qy       159 TTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCC 218

Db       134 TTCGACGGCGCCGCCTAAGCCGGTGAGAATCACCTCCCCAACAGTTGCCGGAAGTTACCC 193

Qy       219 CGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAA 278

Db       194 CGTCGTCTTATTCTTCCATGGATTTTATCTTCGCAACTACTTCTACTCTGATGTTCTTAA 253

Qy       279 CCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCC 338

Db       254 CCACGTCGCTTCTCATGGCTACATTCTTGTAGCCCCACAGTTGTGCAAGATTTTGCCGCC 313

Qy       339 GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCT 398

Db       314 GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAACTGGGCATCGGAAAACCT 373

Qy       399 CAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCA 458

Db       374 CAAAGCTCACCTCCCAAACTCCGTAAACGCTAATGGCAAATACACCGCACTCGTGGGCCA 433

Qy       459 CAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCC 518

Db       434 TAGCCGCGGTGGCAAAACCGCGTTTGCGGTTGCTTTAGGCCATGCCGCAACATTAGACAC 493

Qy       519 ATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACAT 578

Db       494 ATCCTTCAAATTTTCAGCTCTCATAGGAATCGACCCGGTAGCAGGAATCAGCAAATGCAT 553

Qy       579 TAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGT 638

Db       554 GAGAACCGATCCGGAAATCTTAACGTATAAACCGGAATCCTTCGAGATAGACATACCGGT 613

Qy       639 TGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACC 698

Db       614 TGCAGTGGTCGGGACGGGTCTCGGACCCAAGAGTAACGGGATGATGCCACCATGCGCACC 673

Qy       699 AACGG-ACTTAAACCATGAGGAGTTTTACAAAGAGT 733

Db       674 AGCGGAAGTGAACCCATGAGGAGTTTATAAGAGTGT 709


RESULT 9
B20290/c
LOCUS       B20290                   794 bp    DNA     linear   GSS 16-SEP-1997
DEFINITION  T21K8-T7 TAMU Arabidopsis thaliana genomic clone T21K8, genomic
            survey sequence.
ACCESSION   B20290
VERSION     B20290.1  GI:2395344
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
            Ecker,J.
  TITLE     BAC End Sequences at ATGC
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: T21K8-Sp6
            Contact: Ecker J.
            Arabidopsis Thaliana Genome Center
            University of Pennsylvania
            Dept. of Biology, University of Pennsylvania, Philadelphia, PA
            19104
            Tel: 215-898-9384
            Fax: 215-898-8780
            Email: jecker@atgenome.bio.upenn.edu
            Seq primer: T7
            Class: BAC ends
            High quality sequence start: 100
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="T21K8"
                     /sex="hermaphrodite"
                     /clone_lib="TAMU"
                     /note="Vector: BeloBACII; Site_1: HindIII; Site_2:
                     HindIII; Produced by Rod Wing"
ORIGIN

  Query Match           42.6%;  Score 506.6;  DB 9;  Length 794;
  Best Local Similarity   90.0%;  Pred. No. 3.6e-137;
  Matches  561;  Conservative    0;  Mismatches   60;  Indels    2;  Gaps    2;

Qy       358 GACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACT 417

Db       693 GAGAATGTGAAAAGTTGATGAAAGTGGAATTGGGGAAACCTGGAAACTTACNTACCCAAT 634

Qy       418 TCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACG 477

Db       633 TCGGTTAAATGTAATGGNAAATACACTCAACTCGTGGNCCACACCTGCGGTGGGAANAAG 574

Qy       478 GCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCT 537

Db       573 GCGTTTGCGTTTGCGTTAGNCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCT 514

Qy       538 CTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATC 597

Db       513 TTAATAGGGATTGATCCAGTCGCAGGGACTNACAAATACATAAGACAAGATCGCATATTC 454

Qy       598 TTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGA 657

Db       453 TAAACGTATAACCCGGGATCTTTCGAGCTGAACATACCGGTTGCAGTGGT-GGAACCGGA 395

Qy       658 CTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAG 717

Db       394 CTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAA 335

Qy       718 GAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACAT 777

Db       334 GAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACAT 275

Qy       778 ATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGT 837

Db       274 ATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGT 215

Qy       838 AAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCG 897

Db       214 AAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCG 155

Qy       898 TTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCT 957

Db       154 TTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCT 95

Qy       958 TCCG-TTTCTCCGGCCAAGCTTG 979

Db        94 TCCGTTTTATCCGGCCAAGCTTG 72


RESULT 10
AI997409/c
LOCUS       AI997409                 547 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701553812 A. thaliana, Columbia Col-0, root-2 Arabidopsis thaliana
            cDNA clone 701553812, mRNA sequence.
ACCESSION   AI997409
VERSION     AI997409.1  GI:5844314
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D.,
            Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C.,
            Carpio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S.,
            Nobriga,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and
            Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="701553812"
                     /tissue_type="root"
                     /dev_stage="4 - 7 weeks"
                     /clone_lib="A. thaliana, Columbia Col-0, root-2"
                     /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
                     library was derived from untreated root tissue from
                     Arabidopsis thaliana,  Columbia Col-0, at 4 - 7 weeks.
                     Plants were grown in 1:1:1 peat  moss/vermiculite/perlite
                     soil at 22 deg. C +/- 3 deg. C under constant  light, and
                     watered with fertilizer. cDNA synthesis was initiated
                     using a  NotI-oligo(dT) primer. Double-stranded cDNA was
                     blunted, ligated to SalI  adaptors, digested with NotI,
                     size-selected, and cloned into the NotI and  SalI sites of
                     the pSPORT vector."
ORIGIN

  Query Match             41.6%;  Score 493.8;  DB 1;  Length 547;
  Best Local Similarity   94.2%;  Pred. No. 1.9e-133;
  Matches  517;  Conservative    0;  Mismatches   30;  Indels    2;  Gaps    2;

Qy        590 CGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGG 649
              |||||||||||||||||||||||||| |||||||||| |||||||||||| |||| ||||
Db        547 CGCATATCTTAACGTATAAACCGGAA-CTTTCGAGCT-GACATACCGGTTNCAGTNGTGG 490

Qy        650 GAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAA 709
                | |||||||||||||||||| |||| ||||||||||||||||||||||||||||||
Db        489 NACCCGGACTCGGACCGAAGTGNAACANCGTGATGCCACCATGCGCACCAACGGACTTAA 430

Qy        710 ACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATT 769
              |||||||||||| |||||||||||||||||| |||| |||||| |||| ||||||||||
Db        429 ACCATGAGGAGTTNTACAAAGAGTGTAAGGCGNCGAANGCCCATNTCGTNGCTGCGGATT 370

Qy        770 ACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTT 829
              ||| |||||||||||||||||||||||||||||||||| |||||||| |  | ||||||
Db        369 ACGNACATATGGATATGTTGGACGATGATTTGCCCGGTTNTGTTGGGTNTNNGNCCGGTT 310

Qy        830 GTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTG 889
              |||||||||||| |||||||||| |||||||||||||| ||| ||||||||||||||||
Db        309 GTATGTGTAAGANTGGGCAAAGAANAAAGTCTGAGATGANGAGNTTTGTAGGTGGAATTG 250

Qy        890 TGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGA 949
              |||||  | |||||||||| ||||| ||||||||||||||||||||||||||||||||||
Db        249 TGGTTNNGNTTCTCAAGTANAGTTTNTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGA 190

Qy        950 AGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTG 1009
              |||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||
Db        189 AGGATCCTNCCGTTTCTCCGGCCAAGCTTGATCTTTCACCTGAGTTGGAAGAAGCTTCTG 130

Qy        1010 GTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAA 1069
               |||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         129 GTATCNNCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAA 70

Qy        1070 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 1129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          69 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 10

Qy        1130 TTTTTATTA 1138
               |||||||||
Db           9 TTTTTATTA 1


RESULT 11
AV561535
LOCUS       AV561535                413 bp    mRNA    linear   EST 23-FEB-2004
DEFINITION  AV561535 Arabidopsis thaliana green siliques Columbia Arabidopsis
            thaliana cDNA clone SQ153b01F 3', mRNA sequence.
ACCESSION   AV561535
VERSION     AV561535.1  GI:8732961
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7 (3), 175-180 (2000)
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SQ153b01F"
                     /tissue_type="green siliques"
                     /clone_lib="Arabidopsis thaliana green siliques Columbia"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match             34.8%;  Score 413;  DB 1;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 1e-109;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        533 CAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGC 592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGC 60

Qy        593 ATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAA 652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAA 120

Qy        653 CCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACC 712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACC 180

Qy        713 ATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACG 772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACG 240

Qy        773 GACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTA 832
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTA 300

Qy         833 TGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGG 892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGG 360

Qy         893 TTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATT 945
               |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATT 413


RESULT 12
DN779181/c
LOCUS       DN779181                 570 bp    mRNA    linear   EST 05-APR-2005
DEFINITION  E7566 [C4680|ss1_f12] Salinity Subtracted Library Thellungiella
            salsuginea cDNA, mRNA sequence.
ACCESSION   DN779181
VERSION     DN779181.1  GI:62208062
KEYWORDS    EST.
SOURCE      Thellungiella salsuginea
  ORGANISM  Thellungiella salsuginea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Wong,C.E., Li,Y., Whitty,B.R., Diaz-Camino,C., Akhter,S.R.,
            Brandle,J.E., Golding,G.B., Weretilnyk,E.A., Moffatt,B.A. and
            Griffith,M.
  TITLE     Expressed sequence tags from the Yukon ecotype of Thellungiella
            reveal that gene expression in response to cold, drought and
            salinity shows little overlap
  JOURNAL   Plant Mol. Biol. 58 (4), 561-574 (2005)
   PUBMED   16021339
COMMENT     Contact: Moffatt,B
            Department of Biology
            University of Waterloo
            200, University Ave West,Ontario, Canada N2L 3G1
            Tel: (519) 888-4567 ext 2517
            Fax: (519) 746-0614
            Email: moffatt@uwaterloo.ca.
FEATURES             Location/Qualifiers
     source          1..570
                     /organism="Thellungiella salsuginea"
                     /mol_type="mRNA"
                     /cultivar="Yukon"
                     /db_xref="taxon:72664"
                     /tissue_type="Leave"
                     /dev_stage="4-week old plants"
                     /clone_lib="Salinity Subtracted Library"
                     /note="Vector: pCMVSPORT6; Site_1: Notl; Site_2: Notl;
                     Library was prepared following the protocol for BD
                     Clontech PCR-Select cDNA Subtraction Kit and Promega
                     pGEM-T Easy Vector System for cDNA synthesis and cloning"
ORIGIN

  Query Match             34.7%;  Score 412.4;  DB 8;  Length 570;
  Best Local Similarity   85.0%;  Pred. No. 1.6e-109;
  Matches  472;  Conservative    0;  Mismatches   82;  Indels    1;  Gaps    1;

Qy         478 GCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCT 537
               ||||||||||||||  ||||||||||||||||||||||| ||||| |||   ||||||||
Db         570 GCGTTTGCGGTTGCTTTAGGCCATGCCGCAACATTAGACACATCC-TCAAANTTTCAGCT 512

Qy         538 CTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATC 597
               || |||||||| || || || |||||||  | ||||| ||| |||||||||||| | |||
Db         511 CTCATAGGAATCGACCCGGTAGCAGGAATCAGCAAATGCATGAGAACCGATCCGGAAATC 452

Qy         598 TTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGA 657
               |||||||||||||||||||| |||||| | |||||||||||||||||||| || || || 
Db         451 TTAACGTATAAACCGGAATCCTTCGAGATAGACATACCGGTTGCAGTGGTCGGGACGGGT 392

Qy         658 CTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAG 717
               |||||||| ||| | |||    |||||||||||||||||||| ||||  | |||||||||
Db         391 CTCGGACCCAAGAGTAACGGGATGATGCCACCATGCGCACCAGCGGAAGTGAACCATGAG 332

Qy         718 GAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACAT 777
               |||||||| |||||||||||||| || || |  |||||||||||||| ||||| ||||||
Db         331 GAGTTTTATAAAGAGTGTAAGGCCACCAAGGGACATTTCGTGGCTGCTGATTATGGACAT 272

Qy         778 ATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGT 837
               ||||||||||||||||||||||||||||||||||| ||||||||||| ||||||||||||
Db         271 ATGGATATGTTGGACGATGATTTGCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGT 212

Qy         838 AAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCG 897
               ||||| ||| || |||||||   |||||||||||| ||||| |  | ||||||||||||
Db         211 AAGAACGGGAAACGAAAAAATAGTGAGATGAGGAGATTTGTGGGAGGTATTGTGGTTGCA 152

Qy         898 TTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCT 957
               ||||||| ||||||||| |||||| | ||| |||||||||||  |||| |||||||||||
Db         151 TTTCTCAGGTATAGTTTATGGGGTCACAAATCGGAGATTCGAGAGATTCTGAAGGATCCT 92

Qy         958 TCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTC 1017
               || |  |||||||  | |||||||||||  || ||| ||||||| ||||| ||   ||||
Db          91 TCTGCCTCTCCGGTGACGCTTGATCCTTTGCCCGAGCTGGAAGAGGCTTCGGGCTACTTC 32

Qy        1018 GTCTAGATTTGTGTT 1032
               ||||||||| ||| |
Db          31 GTCTAGATTGGTGAT 17


RESULT 13
DN779131/c
LOCUS       DN779131                 547 bp    mRNA    linear   EST 05-APR-2005
DEFINITION  E7515 [C4629|ss1_b09] Salinity Subtracted Library Thellungiella
            salsuginea cDNA, mRNA sequence.
ACCESSION   DN779131
VERSION     DN779131.1  GI:62208012
KEYWORDS    EST.
SOURCE      Thellungiella salsuginea
  ORGANISM  Thellungiella salsuginea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Thellungiella.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Wong,C.E., Li,Y., Whitty,B.R., Diaz-Camino,C., Akhter,S.R.,
            Brandle,J.E., Golding,G.B., Weretilnyk,E.A., Moffatt,B.A. and
            Griffith,M.
  TITLE     Expressed sequence tags from the Yukon ecotype of Thellungiella
            reveal that gene expression in response to cold, drought and
            salinity shows little overlap
  JOURNAL   Plant Mol. Biol. 58 (4), 561-574 (2005)
   PUBMED   16021339
COMMENT     Contact: Moffatt,B
            Department of Biology
            University of Waterloo
            200, University Ave West,Ontario, Canada N2L 3G1
            Tel: (519) 888-4567 ext 2517
            Fax: (519) 746-0614
            Email: moffatt@uwaterloo.ca.
FEATURES             Location/Qualifiers
     source          1..547
                     /organism="Thellungiella salsuginea"
                     /mol_type="mRNA"
                     /cultivar="Yukon"
                     /db_xref="taxon:72664"
                     /tissue_type="Leave"
                     /dev_stage="4-week old plants"
                     /clone_lib="Salinity Subtracted Library"
                     /note="Vector: pCMVSPORT6; Site_1: Notl; Site_2: Notl;
                     Library was prepared following the protocol for BD
                     Clontech PCR-Select cDNA Subtraction Kit and Promega
                     pGEM-T Easy Vector System for cDNA synthesis and cloning"
ORIGIN

  Query Match             33.9%;  Score 403;  DB 8;  Length 547;
```

```
  Best Local Similarity   84.9%;  Pred. No. 9.5e-107;
  Matches  451;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        502 GCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCA 561
              ||||||||||||||| ||||| |||  |||||||||| |||||||| || || || |||
Db        547 GCCGCAACATTAGACACATCCTTCAAATTTTCAGCTCTCATAGGAATCGACCCGGTAGCA 488

Qy        562 GGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTC 621
              ||||  | ||||| ||| ||||||||||| | |||||||||||||||||||||||| |||
Db        487 GGAATCAGCAAATGCATGAGAACCGATCCGGAAATCTTAACGTATAAACCGGAATCCTTC 428

Qy        622 GAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTG 681
              ||| | |||||||||||||||||||| || || || |||||||| ||| | |||    ||
Db        427 GAGATAGACATACCGGTTGCAGTGGTCGGGACGGGTCTCGGACCCAAGAGTAACGGGATG 368

Qy        682 ATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCG 741
              |||||||||||||||||| ||||  | ||||||||||||||| ||||||||||||||
Db        367 ATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATAAAGAGTGTAAGGCC 308

Qy        742 ACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTG 801
              || || |  |||||||||||||| ||||| ||||||||||||||||||||||||||||
Db        307 ACCAAGGGACATTTCGTGGCTGCTGATTATGGACATATGGATATGTTGGACGATGATTTG 248

Qy        802 CCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCT 861
              |||||||||| |||||||||| |||||||||||||||| ||| || |||||||   |
Db        247 CCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACGGGAAACGAAAAAATAGT 188

Qy        862 GAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGT 921
              |||||||||| ||||| || || |||||||||| ||||||| |||||||| ||||||
Db        187 GAGATGAGGAGATTTGTGGGAGGTATTGTGGTTGCATTTCTCAGGTATAGTTTATGGGGT 128

Qy        922 GAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGAT 981
              | ||| ||||||||||  |||| |||||||||||| |  |||||||  | |||||||
Db        127 CACAAATCGGAGATTCGAGAGATTCTGAAGGATCCTTCTGCCTCTCCGGTGACGCTTGAT 68

Qy        982 CCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTT 1032
              ||||  || ||| ||||||| ||||| ||   ||||||||||||| ||| |
Db         67 CCTTTGCCCGAGCTGGAAGAGGCTTCGGGCTACTTCGTCTAGATTGGTGCT 17


RESULT 14
BP605375/c
LOCUS       BP605375                 430 bp    mRNA    linear   EST 25-JUN-2004
DEFINITION  BP605375 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-66-J07 3',
            mRNA sequence.
ACCESSION   BP605375
VERSION     BP605375.1  GI:49252888
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Seki,M., Narusaka,M., Kamiya,A., Ishida,J., Satou,M., Sakurai,T.,
            Nakajima,M., Enju,A., Akiyama,K., Oono,Y., Muramatsu,M.,
            Hayashizaki,Y., Kawai,J., Carninci,P., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A. and Shinozaki,K.
  TITLE     Functional annotation of a full-length Arabidopsis cDNA collection
  JOURNAL   Science 296 (5565), 141-145 (2002)
   PUBMED   11910074
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            reversed clone; Please visit our web site
            (http://pfgweb.gsc.riken.go.jp/) for further details.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL16-66-J07"
                     /lab_host="DH10B"
                     /clone_lib="RAFL16"
                     /note="Site_1: BamHI; Site_2: SalI; dark-grown"
ORIGIN

  Query Match             33.0%;  Score 391.6;  DB 3;  Length 430;
  Best Local Similarity   99.0%;  Pred. No. 2e-103;
  Matches  403;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        766 GATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCC 825
              |||||||||||||| |||||||||||||||||||| ||||||||| ||||||||||||||
Db        407 GATTACGGACATAT-GATATGTTGGACGATGATTTCCCCGGTTTTTTTGGGTTTATGGCC 349

Qy        826 GGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGA 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        348 GGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGA 289

Qy        886 ATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATT 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 ATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATT 229

Qy        946 GTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCT 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 GTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCT 169

Qy       1006 TCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTG 1065
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        168 TCTGGTATCTTCGTCTAGATTTTTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTG 109

Qy       1066 AAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATG 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 AAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATG 49

Qy       1126 TGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCC 1172
              ||||||||||||||||||||||||||||||||||||||||||||||
Db          48 TGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCC 2


RESULT 15
AV817693/c
LOCUS       AV817693                 411 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  AV817693 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-95-F15 3',
            mRNA sequence.
ACCESSION   AV817693
VERSION     AV817693.1  GI:19859588
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 411)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
            and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
```

site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .411
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL09-95-F15"
                     /dev_stage="plants at various developmental stages from
                     germination to mature seeds"
                     /lab_host="DH10B"
                     /clone_lib="RAFL9"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                     hr) treatments"
ORIGIN

  Query Match           32.8%;  Score 389.4;  DB 1;  Length 411;
  Best Local Similarity 99.7%;  Pred. No. 8.8e-103;
  Matches  390;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        786 GTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGG 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 GTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGG 352

Qy        846 GCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAA 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 GCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAA 292

Qy        906 GTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTC 965
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        291 GTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTTTGAAGGATCCTTCCGTTTC 232

Qy        966 TCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGAT 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        231 TCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGAT 172

Qy       1026 TTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTC 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 TTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTC 112

Qy       1086 TAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGA 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        111 TAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGA 52

Qy       1146 TCAAAACATTTGTTATAGTTTTACCCCAAAA 1176
              ||||||||||||||||||||||||||||||
Db          51 TCAAAACATTTGTTATAGTTTTACCCCAAAA 21
```

Search completed: March 20, 2006, 16:52:58
Job time : 5388.1 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:51:15 ; Search time 699.094 Seconds
(without alignments)
11325.590 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1188	100.0	1188	5	AAH41136	Aah41136 Arabidops
2	1188	100.0	1188	12	ADJ98180	Adj98180 Thale cre
3	1115.4	93.9	1157	3	AAC34054	Aac34054 Arabidops
4	975	82.1	975	6	ABZ12353	Abz12353 Arabidops
5	975	82.1	975	6	ADG87791	Adg87791 A. thalia
6	975	82.1	975	6	ADG87590	Adg87590 A. thalia
7	748	63.0	1125	13	ADX51602	Adx51602 Plant ful
8	239.4	20.2	987	6	AAD38074	Aad38074 Grape chl
9	202.6	17.1	1174	6	AAD38077	Aad38077 Soybean c
10	199	16.8	1242	6	AAD38082	Aad38082 Wheat chl
11	194.4	16.4	1104	6	AAD38078	Aad38078 Soybean c
12	188	15.8	1125	6	AAD38079	Aad38079 Soybean c
13	178.8	15.1	1444	6	AAD38080	Aad38080 Tulip chl
14	169.8	14.3	1216	6	AAD38075	Aad38075 Grape chl
15	163.4	13.8	1135	5	AAH41137	Aah41137 Arabidops
16	163.4	13.8	1135	12	ADJ98181	Adj98181 Thale cre
17	151.6	12.8	1340	5	AAH41135	Aah41135 Chenopodi
18	110.8	9.3	1156	6	AAD38086	Aad38086 Wheat chl
19	109.2	9.2	1302	6	AAD38076	Aad38076 Corn chlo
20	106.6	9.0	1244	13	ADX50647	Adx50647 Plant ful
21	106.6	9.0	1244	13	ADX50646	Adx50646 Plant ful
22	98.6	8.3	916	6	AAD38081	Aad38081 Wheat chl
23	82.6	7.0	1274	6	AAD38087	Aad38087 Corn chlo
24	67	5.6	1300	5	AAH41145	Aah41145 Chenopodi
25	63.8	5.4	558	11	ACL32937	Acl32937 Rice abio
26	53	4.5	390	11	ACL27601	Acl27601 Rice abio
27	47	4.0	985	6	ABZ15053	Abz15053 Arabidops
28	45	3.8	1829	2	AAX84982	Aax84982 Human sec
29	45	3.8	1829	8	ACD18908	Acd18908 Novel hum
30	45	3.8	1829	12	ADG78299	Adg78299 Human sec
31	45	3.8	1829	12	ADN60590	Adn60590 Human sec
c 32	44.8	3.8	2000	8	ADA71938	Ada71938 Rice gene
c 33	41.6	3.5	631	5	ADI73251	Adi73251 Human ova
c 34	41.6	3.5	631	5	ADL38382	Adl38382 Human ova
35	41.6	3.5	2000	8	ADA71938	Ada71938 Rice gene
36	40.8	3.4	13376	6	ABL32583	Abl32583 Human imm
c 37	40.4	3.4	6715	13	ADR84259	Adr84259 Aspergill
c 38	39.8	3.4	752	5	ADL44766	Adl44766 Human ova
39	39.6	3.3	604	4	AAH34528	Aah34528 Human col
40	39.6	3.3	1345	12	ADQ22878	Adq22878 Human sof
c 41	39.6	3.3	2000	11	ACL35601	Acl35601 Rice stre
42	39.6	3.3	4419	6	ABS53449	Abs53449 cDNA enco
c 43	39.4	3.3	452	13	ACF83451	Acf83451 Human SIR
c 44	39	3.3	2000	11	ACL35887	Acl35887 Rice stre
45	38.8	3.3	2000	11	ACL35887	Acl35887 Rice stre

ALIGNMENTS

RESULT 1
AAH41136
ID AAH41136 standard; DNA; 1188 BP.
XX
AC AAH41136;
XX
DT 21-AUG-2001 (first entry)
XX
DE Arabidopsis thaliana chlorophyllase coding sequence #1.
XX
KW Chlorophyllase; transgenic plant; ds.
XX
OS Arabidopsis thaliana.
XX
PN JP2001086990-A.
XX
PD 03-APR-2001.
XX
PF 20-SEP-1999; 99JP-00266181.
XX
PR 20-SEP-1999; 99JP-00266181.
XX
PA (KAGO) KAGOME KK.
XX
DR WPI; 2001-338421/36.
DR P-PSDB; AAB99102.
XX
PT DNA encoding chlorophyllase, useful for producing transgenic plants.
XX
PS Claim 3; Page 10-11; 21pp; Japanese.
XX
CC The present sequence is a chlorophyllase protein coding sequence. The
CC chlorophyllase protein coding sequence can be used for the transformation
CC of a plant
XX
SQ Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

Query Match 100.0%; Score 1188; DB 5; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60

```
Db          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60
Qy         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
Db         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
Qy        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
Db        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
Qy        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
Db        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
Qy        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
Db        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
Qy        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
Db        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
Qy        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
Db        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
Qy        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
Db        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
Db        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
Db        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
Db        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
Db        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
Db        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
Db        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
Db        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
Db        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
Db        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
Qy       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
Db       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
Qy       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
Db       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
Qy       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
Db       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
```

RESULT 2
ADJ98180
ID ADJ98180 standard; DNA; 1188 BP.
XX
AC ADJ98180;
XX
DT 06-MAY-2004 (first entry)
XX
DE Thale cress chlorophyllase 1 DNA.
XX
KW phytol kinase; tocopherol biosynthesis; plant; drought resistance;
KW thale cress; ds; chlorophyllase 1.
XX
OS Arabidopsis thaliana.
XX
PN WO2004013312-A2.
XX
PD 12-FEB-2004.
XX
PF 05-AUG-2003; 2003WO-US025276.
XX
PR 05-AUG-2002; 2002US-0400689P.
PR 05-AUG-2003; 2003US-00634548.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Norris SR, Lincoln K, Abad MS, Eilers R, Hartsuyker KK;
PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;
PI Venkatesh TV;
XX
DR WPI; 2004-157125/15.
XX
PT New phytol kinase polynucleotides, useful in mediating tocopherol
PT biosynthesis and in producing plants with increased drought resistance.
XX
PS Claim 23; SEQ ID NO 18; 189pp; English.
XX
CC The invention relates to a novel substantially purified nucleic acid
CC molecule encoding a phytol kinase. The nucleic acid molecules and
CC polypeptides of the invention may be useful in mediating tocopherol
CC biosynthesis and in producing plants with increased drought resistance.
CC The current sequence is that of the thale cress chlorophyllase 1 DNA of
CC the invention.
XX
SQ Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

```
  Query Match             100.0%;  Score 1188;  DB 12;  Length 1188;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60
Db          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60
Qy         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
Db         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
Qy        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
Db        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
Qy        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
Db        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
```

```
Qy        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300

Qy        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360

Qy        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420

Qy        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480

Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540

Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600

Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660

Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720

Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780

Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840

Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900

Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960

Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020

Qy       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080

Qy       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140

Qy       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188


RESULT 3
AAC34054
ID   AAC34054 standard; DNA; 1157 BP.
XX
AC   AAC34054;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 5284.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
```

PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 93.9%; Score 1115.4; DB 3; Length 1157;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1151; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

```
Qy          2 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAG 64

Qy         62 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 AGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCC 124

Qy        122 CGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 181
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        125 CGACAACCGAGATACCTGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGG 184

Qy        182 TGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 241
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        185 TGAGAATCACCTGTCCA--AGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCT 242

Qy        242 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACA 302
```

```
Qy        302 TTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 361
              ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TTCTTGTAGCCC--CAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACG 360

Qy        362 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGG 420

Qy        422 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGT 480

Qy        482 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAA 540

Qy        542 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAA 600

Qy        602 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCG 660

Qy        662 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGT 720

Qy        722 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGG 780

Qy        782 ATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGA 841
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        781 ATATGTTGGACGATGATTTGCCCGGTTTTGTT-GGTTTATGGCCGGTTGTATGTGTAAGA 839

Qy        842 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 ATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTC 899

Qy        902 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 TCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCG 959

Qy        962 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCT 1019

Qy       1022 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 AGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGT 1079

Qy       1082 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 TTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAAC 1139

Qy       1142 TCGATCAAAACATTTGT 1158
              |||||||||||||||||
Db       1140 TCGATCAAAACATTTGT 1156
```

RESULT 4
ABZ12353
ID ABZ12353 standard; DNA; 975 BP.
XX
AC ABZ12353;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 158.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 158; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

Query Match 82.1%; Score 975; DB 6; Length 975;
Best Local Similarity 100.0%; Pred. No. 3.9e-285;
Matches 975; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023
              |||||||||||||||
Db        961 GGTATCTTCGTCTAG 975
```

```
RESULT 5
ADG87791
ID   ADG87791 standard; cDNA; 975 BP.
XX
AC   ADG87791;
XX
DT   22-APR-2004  (first entry)
XX
DE   A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #233.
XX
KW   Pathogen infection-related gene; plant; Peronospora parasitica;
KW   defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW   oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200222675-A2.
XX
PD   21-MAR-2002.
XX
PF   14-SEP-2001; 2001WO-US028506.
XX
PR   15-SEP-2000; 2000US-0232778P.
PR   22-JUN-2001; 2001US-0300183P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (UYNC-) UNIV NORTH CAROLINA.
PA   (GLAZ/) GLAZEBROOK J.
PA   (WANG/) WANG X.
PA   (DANG/) DANGL J L.
PA   (EULG/) EULGEM T.
PA   (ZHUT/) ZHU T.
XX
PI   Glazebrook J,  Wang X,  Dangl JL,  Eulgem T,  Zhu T;
XX
DR   WPI; 2002-292409/33.
XX
PT   Novel isolated polynucleotide, useful for conveying pathogen resistance
PT   to plants, and for identifying plants infected with a pathogen.
XX
PS   Claim 3; SEQ ID NO 233; 605pp; English.
XX
CC   The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC   ADG87557}) whose expression is altered in response to pathogen infection,
CC   and to homologues of these genes from other plants or fungi, especially
CC   from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC   cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC   expression of genes of the invention was upregulated or downregulated in
CC   Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC   indicating that they play a role in defence mechanisms. The genes of the
CC   invention are regulated by RPP7 or RRP8 which act via unconventional
CC   signalling cascades, or by the RPP4-dependent pathway. The invention also
CC   relates to polypeptides encoded by the pathogen infection-related genes;
CC   promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC   ; expression cassettes, host cells and pathogen-resistant transgenic
CC   plants and their progeny comprising a polynucleotide of the invention;
CC   and a method of identifying a plant cell infected with a pathogen. The
CC   polynucleotide sequences and methods of the invention are useful for
CC   identifying plants infected with a pathogen, and for conferring
CC   resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC   neamotodes and insects (e.g., aphids). The present sequence represents an
CC   Arabidopsis thaliana gene whose expression is altered in response to
CC   Peronospora parasitica infection. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

  Query Match             82.1%;  Score 975;  DB 6;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-285;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023
              ||||||||||||||||
Db        961 GGTATCTTCGTCTAG 975
```

```
RESULT 6
ADG87590
ID   ADG87590 standard; cDNA; 975 BP.
XX
AC   ADG87590;
XX
DT   22-APR-2004  (first entry)
XX
DE   A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #32.
XX
KW   Pathogen infection-related gene; plant; Peronospora parasitica;
KW   defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW   oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200222675-A2.
XX
PD   21-MAR-2002.
XX
PF   14-SEP-2001; 2001WO-US028506.
XX
PR   15-SEP-2000; 2000US-0232778P.
PR   22-JUN-2001; 2001US-0300183P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (UYNC-) UNIV NORTH CAROLINA.
PA   (GLAZ/) GLAZEBROOK J.
PA   (WANG/) WANG X.
PA   (DANG/) DANGL J L.
PA   (EULG/) EULGEM T.
PA   (ZHUT/) ZHU T.
XX
PI   Glazebrook J,  Wang X,  Dangl JL,  Eulgem T,  Zhu T;
XX
DR   WPI; 2002-292409/33.
XX
PT   Novel isolated polynucleotide, useful for conveying pathogen resistance
PT   to plants, and for identifying plants infected with a pathogen.
XX
PS   Claim 3; SEQ ID NO 32; 605pp; English.
XX
CC   The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC   ADG87557}) whose expression is altered in response to pathogen infection,
CC   and to homologues of these genes from other plants or fungi, especially
CC   from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC   cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC   expression of genes of the invention was upregulated or downregulated in
CC   Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC   indicating that they play a role in defence mechanisms. The genes of the
CC   invention are regulated by RPP7 or RRP8 which act via unconventional
CC   signalling cascades, or by the RPP4-dependent pathway. The invention also
CC   relates to polypeptides encoded by the pathogen infection-related genes;
CC   promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC   ; expression cassettes, host cells and pathogen-resistant transgenic
CC   plants and their progeny comprising a polynucleotide of the invention;
CC   and a method of identifying a plant cell infected with a pathogen. The
CC   polynucleotide sequences and methods of the invention are useful for
CC   identifying plants infected with a pathogen, and for conferring
CC   resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC   neamotodes and insects (e.g., aphids). The present sequence represents an
CC   Arabidopsis thaliana gene whose expression is altered in response to
CC   Peronospora parasitica infection. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;
```

```
  Query Match             82.1%;  Score 975;  DB 6;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-285;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy         109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy         169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy         229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy         289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy         349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy         409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy         469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy         529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy         589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023
              |||||||||||||||
Db        961 GGTATCTTCGTCTAG 975
```

```
RESULT 7
ADX51602
ID   ADX51602 standard; cDNA; 1125 BP.
XX
AC   ADX51602;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 26342.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 26342; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This sequence represents a plant full length insert
CC   polynucleotide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 1125 BP; 293 A; 260 C; 274 G; 298 T; 0 U; 0 Other;

  Query Match           63.0%;  Score 748;  DB 13;  Length 1125;
  Best Local Similarity   82.0%;  Pred. No. 4e-216;
  Matches  873;  Conservative    0;  Mismatches  190;  Indels    1;  Gaps    1;

Qy         21 AACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTC 80
              |||  |  | || || | | | |||||||||||||| ||  |||||||||   |||||||
Db         11 AACACACTCTTTTTCCAATATTATACAAATGGCGGGGAAGGAGGACAGTGAGACGTTTTA 70

Qy         81 CTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGT 140
              ||| | || ||||||   ||| |||||| |||||||||| || ||||||| ||| || |
Db         71 CTCGGCGGCAACTCCTTTGGCGTTTGAGTTAGGCAGCCTTCCAACAACCGTGATCCCCGC 130

Qy        141 GGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAAC 200
              || |||  || ||  ||||  || ||||| ||||| || || | ||||||||  |||||
Db        131 AGACCCGTCGGCAACCGATTTGACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAAC 190

Qy        201 AGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTT 260
              |||||||||||||| |||||||| |||||||||||||| || |||||||| ||||||||
Db        191 CGTCGCCGGAACTTACCCCGTCGTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTT 250

Qy        261 CTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTT 320
              |||||||| ||| |||||||| | ||||| ||||| |||||| |||||||||||||| |
Db        251 CTACTCTGATGTTATTAACCACGTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCT 310

Qy        321 GTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAA 380
              ||||||  | |||||||||||||||||||||||||||||||||||||||  ||||||||
Db        311 GTGCAAGATTTTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAA 370

Qy        381 CTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATA 440
              |||| | ||| |||||||||||||||||| |||| || ||||| |||||||| || ||
Db        371 CTGGACTTCGAAAAACCTCAAAGCTCACCTCCCAAGCTCAGTAAACGCTAATGGCAACTA 430

Qy        441 CACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCA 500
              |||| |||||||||||| |||||||||| ||||| |||||||||||||| |||||||
Db        431 CACCGCACTCGTGGGCCATAGCCGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCA 490

Qy        501 TGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGC 560
              ||||||||| |||||||||||||| |||||||||||  ||||||| ||||||| ||
Db        491 CGCCGCAACACTAGACCCATCCATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGC 550

Qy        561 AGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTT 620
              |||||  | ||||| ||| |||||||||||  | |||||||||||| |||||||||| ||
Db        551 AGGAATCAGCAAATGCATAAGAACCGATCCCGAAATCTTAACGTACAAACCGGAATCATT 610

Qy        621 CGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGT 680
              ||| |||||||| |||||||||||| | || || || |||||||||||| | ||||   |
```

```
Db        611 CGACCTGGACATGCCGGTTGCAGTGATCGGTACGGGTCTCGGACCGAAGAGTAACATGCT 670

Qy        681 GATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGC 740
              ||||||||||||||||||| ||||  | |||||||||||||||||  |  ||||||||||
Db        671 GATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAGGC 730

Qy        741 GACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTT 800
              || || |  ||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        731 TACCAAGGGGCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATAATTT 790

Qy        801 GCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTC 860
              |||||||||||  |||||||||||  ||||||||||||||||||  |  || | ||||||
Db        791 GCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACAGTAAACGCAAAAAGAG 850

Qy        861 TGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGG 920
              ||||||||| |||||||| |||||||||||||||||||||||| |||||||||  | | |||
Db        851 TGAGATGAGAAGCTTTGTTGGTGGAATTGTGGTTGCGTTTCTAAAGTATAGTCT-TTGGG 909

Qy        921 TGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGA 980
              ||||||  | |||||||||  |||||||||||||||| || ||||||||||  | ||||||
Db        910 TGAAAAGTCAGAGATTCGACAGATTGTGAAGGATCCCTCTGTTTCTCCGGCGAGGCTTGA 969

Qy        981 TCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTA 1040
              |||||  || ||| ||||||| |||||||||    |||||||||||||| |
Db        970 TCCTTTGCCGGAGCTGGAAGAGGCTTCTGGTTATCTCGTCTAGATTTGTGGTTAAGGAAT 1029

Qy       1041 TTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTT 1084
              |||  |    ||  | |   | |||  | ||  ||  || |
Db       1030 GTATTGGTATCCTCAGGGAAGCTCCAAACTCGTAATAAACTTAT 1073
```

RESULT 8
AAD38074
ID AAD38074 standard; cDNA; 987 BP.
XX
AC AAD38074;
XX
DT 10-SEP-2002 (first entry)
XX
DE Grape chlorophyllase cDNA from vdb1c.pk002.p19:fis clone.
XX
KW Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
OS Vitis sp.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .858
FT		/*tag= a
FT		/product= "Grape chlorophyllase protein from vdb1c.pk002.p19:fis clone"
FT		/EC_number= "3.1.1.14"
FT		/note= "CDS does not include start codon"
FT		/partial

XX
PN WO200229022-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031059.
XX
PR 05-OCT-2000; 2000US-0238161P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Thorpe CJ;
XX
DR WPI; 2002-444102/47.
DR P-PSDB; AAE23776.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 47; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is grape
CC chlorophyllase cDNA from vdb1c.pk002.p19:fis clone
XX
SQ Sequence 987 BP; 260 A; 224 C; 218 G; 285 T; 0 U; 0 Other;

Query Match 20.2%; Score 239.4; DB 6; Length 987;
Best Local Similarity 57.6%; Pred. No. 1.3e-61;
Matches 473; Conservative 0; Mismatches 336; Indels 12; Gaps 2;

```
Qy        167 CACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTT 226
              | || || || ||  ||  ||       |||||  |    || || || || ||  | |
Db         26 CCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTCTCT 85

Qy        227 TATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCG 286
              | ||  | ||||||||  | || ||||||  |||||||| ||  |  | ||| | | |||
Db         86 TGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCATTT 145

Qy        287 CTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGC 346
              |||| ||||| | |||| | || || || ||||| | |  | |||| || ||     ||
Db        146 CTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTGGAA 205

Qy        347 AAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTC 406
              ||  |  |   || | |  || | ||| |||  |||   |  |||  || ||
Db        206 TTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAATCTG 265

Qy        407 ACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCG 466
              || |||       || ||  |  |    |      | ||    |||||||| | |
Db        266 TGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCAGAG 325

Qy        467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCA 526
              | || || || || ||||| | || |||||  ||||          ||  ||||| |||
Db        326 GGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGC---------TGATACATCCCTCA 376

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
              || ||||| ||| |||||  | || || || |  ||     |  ||||        ||| |
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        587 ATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGG 646
              |||  | ||| |||| |||   ||  | || ||| | || |  || || |||   ||
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAGCAATCCCAGTTTGCGTAA 496

Qy        647 TGGGAACCGGACTCGGACCGAAG---TGGAACAACGTGATGCCACCATGCGCACCAACGG 703
              | || || ||  | ||     ||   |||||  | | |     |||| || |||    |
Db        497 TCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAGATG 556

Qy        704 ACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTG 763
              | |||||||  |||||||  ||  ||||||||  |  |      |||| || ||| ||
Db        557 GAGTGAACCATGTAGAGTTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGACTA 616

Qy        764 CGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGG 823
              | || || || ||  |||| ||||| || ||| || |  | || |   |||||   ||
Db        617 CTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGATTT 676

Qy        824 CCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTG 883
              | |||| ||| || |||| |||| |  |    | |      ||||||||  | ||| ||||
Db        677 CGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGGGTG 736

Qy        884 GAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGA 943
              |  || | ||||| ||  | ||| |  ||||   ||| | |  |  ||  |    |  |
Db        737 GCCTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATTTCAAAGCCA 796

Qy        944 TTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCT 984
              ||||  | || |||    |  |||| |  ||||||||||||
```

```
Db        797 TTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCT 837


RESULT 9
AAD38077
ID   AAD38077 standard; cDNA; 1174 BP.
XX
AC   AAD38077;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Soybean chlorophyllase cDNA from sfl1.pk0046.f8 clone.
XX
KW   Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             12. .959
FT                   /*tag=  a
FT                   /product= "Soybean chlorophyllase protein from
FT                   sfl1.pk0046.f8 clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23779.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 52; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is soybean
CC   chlorophyllase cDNA from sfl1.pk0046.f8 clone
XX
SQ   Sequence 1174 BP; 318 A; 226 C; 260 G; 370 T; 0 U; 0 Other;

  Query Match             17.1%;  Score 202.6;  DB 6;  Length 1174;
  Best Local Similarity   56.9%;  Pred. No. 2.2e-50;
  Matches  480;  Conservative    0;  Mismatches  339;  Indels   24;  Gaps    5;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
              |||| ||| |  |||| |||||||| ||  ||||     |||| ||  | ||  |
Db        102 AATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTGGCTCA 161

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
              || || ||  | || |||| |||||| |||| ||||||||   || |||||||||  |
Db        162 TACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTGAGCTC 221

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTG 333
              ||   |||||| ||||| ||||| | |||  |||| || || ||| ||||        |
Db        222 CTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGAGTGTA 281

Qy        334 CCGCC------GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGG--- 384
              | | |      || |     | ||||| |  ||| || |  ||  |  | |||
Db        282 CGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATTGGCTA 341

Qy        385 GCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACC 444
              ||   |||    ||  || |||  || |||       ||  | || ||      |  |
Db        342 GCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATAAATTG 401

Qy        445 TCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCC 504
                  |    || ||||||   ||||| ||||| |  ||||| || || || || ||||
Db        402 GTTTTATCAGGTCACAGCAAGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTTATGCT 461

Qy        505 GCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGA 564
                ||       |   | ||| ||||||||| ||| |||| || || || || || ||
Db        462 AAAA---------CTAACCTCAAGTTTTCAGCACTAGTAGGCATAGACCCTGTGGCTGGC 512

Qy        565 ACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAG 624
              |    ||||        |||   ||| | |||  | ||    |   |  |||| ||| |
Db        513 CCATGTAAATCTTGCGAAACATTTCCTCCTATTCTCACTGGCATGTCCCAATCCTTCAAT 572

Qy        625 CTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTG 681
              || |||||||  |||  ||  | || || || || || ||    ||||   ||     |
Db        573 TTGAACATACCCATTGTTGTAATTGGCACTGGGCTAGGCCCAGAGAAGGCTAATTTTTTT 632

Qy        682 ATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCG 741
              || |||||||  || ||    |   | ||||| |||||||| |||  |||| ||  |
Db        633 ATTCCACCATGTGCTCCTGATGGGGTGAACCATAAGGAGTTTTTCAATAAGTGCAAACCC 692

Qy        742 ACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTG 801
              |    || ||||| || ||  | || || || || ||||| ||||||||  ||||
Db        693 CCTTGTGCACATTTTGTTGCAACTGAGTATGGTCACATGGACATGTTGGATGATGTGACA 752

Qy        802 CCCGGTTTTGTTGGGT---TTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAG 858
              || || ||  ||||||    |  || |   |||||| || ||| ||||| |  |    | |
Db        753 CCTGGCTTAATTGGGTCAATATTGTCAAATTGTATATGCAAGGATGGGAAGGGTCCTAGG 812

Qy        859 TCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGG 918
                  |||||| ||     || || ||  | || || || ||  | | |       |||
Db        813 GACTTGATGAGAAGGACCGTGGGAGGGTTGGTTGTGGCCTTCTTAAGGGCACAGTTGAAT 872

Qy        919 GGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTT 978
              ||   |     ||| ||       || ||  | |||||    || ||||  | || ||
Db        873 GGCCTATGGAAGGATTTTAATGCTGTTTTGGCGAATCCTAATCTTGCTCCTACTAAACTG 932

Qy        979 GAT 981
              |||
Db        933 GAT 935


RESULT 10
AAD38082
ID   AAD38082 standard; cDNA; 1242 BP.
XX
AC   AAD38082;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Wheat chlorophyllase cDNA from wle1n.pk0058.a4:fis clone.
XX
KW   Wheat; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   CDS             47. .1006
FT                   /*tag=  a
FT                   /product= "Wheat chlorophyllase protein from
FT                   wle1n.pk0058.a4:fis clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
```

XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031059.
XX
PR 05-OCT-2000; 2000US-0238161P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Thorpe CJ;
XX
DR WPI; 2002-444102/47.
DR P-PSDB; AAE23784.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 59; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is wheat
CC chlorophyllase cDNA from wle1n.pk0058.a4:fis clone
XX
SQ Sequence 1242 BP; 266 A; 382 C; 343 G; 251 T; 0 U; 0 Other;

Query Match 16.8%; Score 199; DB 6; Length 1242;
Best Local Similarity 55.8%; Pred. No. 2.8e-49;
Matches 476; Conservative 0; Mismatches 350; Indels 27; Gaps 4;

```
Qy        137 CGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTC 196
              || |  |   ||  || ||||   |  |  |||| || |  ||||||   |||  |   |
Db        153 CGATTCAGGTGGATGAGAATGCGGCGCCGACACCACCGATCCCGGTGCTGATCGTCGCAC 212

Qy        197 CAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACT 256
              | |  |  || ||||| || ||||| |   |  |||| || |||||  |  || |  |||
Db        213 CCAAGGATGCAGGAACCTACCCCGTGGCCATGCTCTTGCACGGCTTCTTCCTCCATAACC 272

Qy        257 ACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCAC 316
              ||||||||    || ||||    ||| |||| || || || | |||  |||| || || |
Db        273 ACTTCTACGAACACCTTCTCCGGCACGTCGCATCCCACGGCTTCATCATTGTCGCGCCCC 332

Qy        317 AGTTGTG---CAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTG 373
              ||||  |   ||   |  | ||  |||| |    || |||  | | ||  || |  |  |
Db        333 AGTTCAGCATCAGTATCATACCTTCGGGTGACGCAGAGGACATCGCCGCGGCAGCCAAGG 392

Qy        374 TGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATG 433
              ||  | |||||    | ||   ||||    |    || || |     ||  |  |  |
Db        393 TGGCAGACTGGCTCCCCGACGGCCTCCCGTCCGTGCTGCCCAAAGGCGTCGAGCCGGAGC 452

Qy        434 GAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGC 493
                  |   | |  | |  |||||||||||| || ||  | ||||| ||  |  | ||
Db        453 TCTCGAAGCTCGCCTTGGCCGGCCACAGCCGAGGAGGCCACACGGCTTTCTCCCTGGCCT 512

Qy        494 TAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATC 553
              | || || |||          |  |     | || || || || || || || ||| | || |
Db        513 TGGGGCACGCC---------AAGACCCAGCTAACCTTCTCCGCGCTCATCGGACTCGACC 563

Qy        554 CAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGG 613
              | ||||| || ||    || | |         | | ||  | ||| | || ||  | |||
Db        564 CCGTCGCCGGCACGGGGAAGTCCTCCCAGCTCCAGCCCAAGATCCTCACCTACGAGCCGT 623

Qy        614 AATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGA 673
                || ||||   |||  || |||||    ||  | || |||||  |||||   | ||  ||
Db        624 CCTCCTTCGGCATGGCGATGCCGGTGCTGGTCATCGGCACCGGGCTCGGCGAGGAGAAGA 683

Qy        674 ACAACGTGATG---CCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAG 730
```

```
              | |||  |   |    || || |||||||| | |||| | |||||  |  |||||  |||   |
Db        684 AGAACATATTCTTCCCTCCCTGCGCACCCAAGGACGTGAACCACGCCGAGTTCTACCGCG 743

Qy        731 AGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGG 790
              |||| | | || |      | | || ||| |   ||| ||||| ||| |||| ||| |||
Db        744 AGTGCAGGCCGCCCTGCTACTACTTTGTGACCAAGGACTACGGGCATCTGGACATGCTGG 803

Qy        791 ACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA 850
              |||| ||    |||   || |             |   ||  | || ||| |||||  |
Db        804 ACGACGACGCCCCCAAGTTCAT------------CACCTGCGTCTGCAAGGATGGGAACG 851

Qy        851 GAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATA 910
              |    |||    ||||| |||| |  || | ||| ||  |||| || ||||| ||   |
Db        852 GGTGCAAGGGCAAGATGCGGAGGTGCGTTGCTGGGATCATGGTGGCATTTCTTAATGCTG 911

Qy        911 GTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGG 970
              |||  |  || | || ||  ||       ||  |||  || ||  | ||| | ||
Db        912 CCTTGGGTGAGAAAGATGCAGATCTTGAGGCCATACTGAGAGACCCGGCGGTTGCACCCA 971

Qy        971 CCAAGCTTGATCC 983
              ||| ||||||||||
Db        972 CCACGCTTGATCC 984
```

RESULT 11
AAD38078
ID AAD38078 standard; cDNA; 1104 BP.
XX
AC AAD38078;
XX
DT 10-SEP-2002 (first entry)
XX
DE Soybean chlorophyllase cDNA from sfl1n1.pk002.m10:fis clone.
XX
KW Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 26. .967
FT /*tag= a
FT /product= "Soybean chlorophyllase protein from
FT sfl1n1.pk002.m10:fis clone"
FT /EC_number= "3.1.1.14"
XX
PN WO200229022-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031059.
XX
PR 05-OCT-2000; 2000US-0238161P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Thorpe CJ;
XX
DR WPI; 2002-444102/47.
DR P-PSDB; AAE23780.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 53; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced

```
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is soybean
CC   chlorophyllase cDNA from sfl1n1.pk002.m10:fis clone
XX
SQ   Sequence 1104 BP; 308 A; 225 C; 249 G; 322 T; 0 U; 0 Other;

  Query Match             16.4%;  Score 194.4;  DB 6;  Length 1104;
  Best Local Similarity   56.3%;  Pred. No. 6.6e-48;
  Matches  473;  Conservative    0;  Mismatches  346;  Indels   21;  Gaps    5;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
              | || |||     |||| ||||||||  ||  |||       ||||| ||  | ||  |
Db        113 AGTGCTTCCTTCTCACCTCCAAAACCATTGCTAATTTTTACACCAACTGTTCCTGGCGCA 172

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
              || || ||  | || ||| ||||||||||||   ||||||| | || ||||||  |  |
Db        173 TACCCTGTAATATTGTTCGTCCATGGCTTTTTCATTCGCAATTTCTACTACTCAAAGCTC 232

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCC---ACAGTTGTGCAAATTA 330
              ||    |||||| |  || ||||| | |||  | || || ||   || ||| | |||
Db        233 CTAGCCCACATAGTCTCACATGGATTCATAATCGTTGCTCCTCAACTGTTTTCCAATGGG 292

Qy        331 TTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCG 390
              | ||   |   || |      |||||||  |||| ||||  ||      | |||  | |
Db        293 CTTCCTATGTATGGACCCACTGAAGTGGAATATGCAGGAAAAGTTGCGGATTGGATAGCT 352

Qy        391 GAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTC 450
              ||  | ||  ||  |   || |||       ||  | || |||     |  |        ||
Db        353 GAGGAGCTTCAACATTTGCTTCCAGAGAACGTTGAAGCCAATTTGGACAAACTGGTTCTA 412

Qy        451 GTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACA 510
                || |||||  | ||||||||||| |  ||||| || || || || ||| |||||
Db        413 TCAGGTCACAGTAGGGGTGGGAAAACTGTATTTGCTGTGGCTCTTGGTCATGC------- 465

Qy        511 TTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAAC 570
                | |  |     ||| |||||||| ||  |||| || || || || || || || ||
Db        466 --AAAAACTAATCTCAAGTTTTCAGCACTTGTAGGCATAGACCCTGTGGCTGGCACATCT 523

Qy        571 AAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGAC 630
              |||||   ||||||   ||| |||||  | ||    || ||    || || ||  || |
Db        524 AAATATTGTAGAACACGTCCTCATATTCTCACTGGCAAGCCACGGTCCTTTGATTTGAAA 583

Qy        631 ATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTGATGCCA 687
              || || |||| |||  | || || ||| | || ||   ||||   ||        |  |||
Db        584 ATGCCAGTTGAAGTAATTGGCACTGGATTGGGCCCAGAGAAGCTTAATTGTTGTACTCCA 643

Qy        688 CCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAA 747
              || || || ||    |  | ||| || ||||||| | ||| ||||| ||  |  |
Db        644 CCGTGTGCTCCTGATGGGGTGAACTATAAGGAGTTCTTCAACGAGTGCAAACCCCCTTGT 703

Qy        748 GCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGT 807
              ||  | || || |  ||  | || || || ||||| |||||| | |||||    || ||
Db        704 GCTAAATTTGTTGTAGCAAAGTATGGTCACATGGACATGTTGAATGATGACACACCAGGG 763

Qy        808 TTTGTTGGG---TTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA---GAAAAAAGTCT 861
              |  |||||   ||  || |    ||||||||||||||||||    |   |     | |
Db        764 CTAATTGGGACATTGGTGTCAAAGTGTATGTGTAAGAATGGGACGACGGGTCCTAGGGAC 823

Qy        862 GAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGT 921
              |||||| ||        || ||  | || ||  | ||  | | |       |||   |
Db        824 TTGATGAGAAGGACCACTGGAGGGTTGGTTGTGTCCTTCTTGAGGGCACAATTGAATGAC 883

Qy        922 GAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGAT 981
              |     |||  ||      ||| | ||||| |||    || ||||  |  |  | |||
Db        884 CTATGGAAGGATTTTGATGCTATTTTAAAGGACCCTAATCTTGCTCCCACTGAAGTCGAT 943


RESULT 12
AAD38079
ID   AAD38079 standard; cDNA; 1125 BP.
XX
AC   AAD38079;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Soybean chlorophyllase cDNA from sl2.pk130.f15:fis clone.
XX
KW   Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             42. .992
FT                   /*tag=  a
FT                   /product= "Soybean chlorophyllase protein from
FT                   sl2.pk130.f15:fis clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23781.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 55; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is soybean
CC   chlorophyllase cDNA from sl2.pk130.f15:fis clone
XX
SQ   Sequence 1125 BP; 315 A; 219 C; 262 G; 329 T; 0 U; 0 Other;

  Query Match             15.8%;  Score 188;  DB 6;  Length 1125;
  Best Local Similarity   56.6%;  Pred. No. 5.8e-46;
  Matches  479;  Conservative    0;  Mismatches  340;  Indels   27;  Gaps    6;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
              | || |||  |  |||| ||||||||  ||   ||       ||||| ||  | ||
Db        132 ACTGCTTCTTCCTCACCTCCAAAACCATTGCTTATTTTTACACCAACCGTGCCTGGCTTA 191

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
              || || ||  | || || | |||||||||| | ||||||    || |||||||  |  |
Db        192 TACCCTGTAATATTGTTTTGCCATGGCTTTTGTATTCGCACTAGCTACTACTCTAAGCTC 251

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACA---GTTGTGCAAATTA 330
              ||    |||||| | ||| ||||| | ||| ||||| || || ||    ||| | ||
Db        252 CTAGCCCACATAGTTTCACATGGATTCATACTTGTTGCTCCTCAGCTGTTTTCCATTGGG 311

Qy        331 TTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCG 390
              ||||   |   || | ||  |||||  |   ||  ||||| ||  |  | |||  |
Db        312 GTGCCTATGTTTGGACCAGAAGAAGTTAAGTGTGAAGGAAGAGTTGTGGATTGGCTAGAT 371

Qy        391 GAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTC 450
              |    ||  || |    || ||        ||  | || ||     |  |        ||
Db        372 AACGGGCTTCAACCATTGCTTCCCGAGAGCGTTGAAGCCAAACTGGAGAAACTGGTTCTA 431
```

```
Qy        451 GTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACA 510
              || || ||||||   ||||| ||||| || ||||| || || || ||  |
Db        432 GTAGGTCACAGCAAGGGTGGAAAAACAGCATTTGCTGTGGCACTTGGTTA--------C 482

Qy        511 TTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGG------A 564
              |  |  ||    ||| ||||||||| || ||||| || ||||| || || ||      |
Db        483 TGTAAAACAAAGCTCAAGTTTTCAGCACTCATAGGCATAGATCCTGTGGCTGGCGTATCA 542

Qy        565 ACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAG 624
              |      ||       | || |   |||  ||||| | ||    |   ||     || ||  |
Db        543 AAGTGTAAGCCTTGTCGATCACTTCCTGATATCCTCACAGGTGTGCCACGGTCCTTTAAT 602

Qy        625 CTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTG 681
              ||| |||||| ||||| ||  | ||||| ||  | || ||   ||||   ||      |
Db        603 CTGAACATACCTGTTGCTGTAATTGGAACTGGGTTGGGCCCAGAGAAGGCTAATTCTCTT 662

Qy        682 ATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCG 741
              | |||||||| || ||||   |    | |||||| | ||||||| |   ||||| ||  |
Db        663 TTTCCACCATGTGCTCCAAATGGGGTGAACCATAAAGAGTTTTTCTCTGAGTGCAAACCA 722

Qy        742 ACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTG 801
              | |  ||  |||| || ||  ||||||| || || ||||| |||||||| |||||
Db        723 CCTAGTGCATATTTTGTTGCAACGGATTATGGTCACATGGACATGTTGGATGATGAAACA 782

Qy        802 CCCGGTTTTGTTGGG---TTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA---GAAAA 855
              || ||  | |||||    | ||| |    |||||||| |||||||| | |    |
Db        783 CCAGGGGTAATTGGGACAATGATGTCAAAGTGTATGTGCAAGAATGGGAAGAAGGGTCCT 842

Qy        856 AAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTG 915
              | |     |||||| ||   ||| || ||  | || || || ||  | |        |||
Db        843 AGGGACTTGATGAGAAGGACTGTGGGAGGGTTGGTTGTGGCCTTCTTGAGAGCACAGTTG 902

Qy        916 TGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAG 975
                |   |      |||  ||      ||| |   |  ||||    |  |||| |||||
Db        903 AATGAGCAGTGGAAGGATTTTGATGCTATTTTAGCGAGTCCTAATCTAGCTCCCGCCAAA 962

Qy        976 CTTGAT 981
              ||||||
Db        963 CTTGAT 968
```

RESULT 13
AAD38080

```
ID   AAD38080 standard; cDNA; 1444 BP.
XX
AC   AAD38080;
XX
DT   07-AUG-2003  (revised)
DT   10-SEP-2002  (first entry)
XX
DE   Tulip chlorophyllase cDNA from etp1c.pk005.d16:fis clone.
XX
KW   Tulip; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Tulipa gesneriana.
XX
FH   Key             Location/Qualifiers
FT   CDS             97. .1107
FT                   /*tag= a
FT                   /product= "Tulip chlorophyllase protein from
FT                   etp1c.pk005.d16:fis clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23782.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 56-57; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is tulip
CC   chlorophyllase cDNA from etp1c.pk005.d16:fis clone. (Updated on 07-AUG-
CC   2003 to correct OS field.)
XX
SQ   Sequence 1444 BP; 460 A; 295 C; 298 G; 391 T; 0 U; 0 Other;
```

Query Match 15.1%; Score 178.8; DB 6; Length 1444;
Best Local Similarity 54.1%; Pred. No. 4.1e-43;
Matches 438; Conservative 0; Mismatches 357; Indels 15; Gaps 3;

```
Qy        156 TGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTA 215
              ||  ||  || |||| |  || ||        ||| |   ||||   |  |  ||    ||
Db        195 TGTCTCCGCACCACCTCGGAAGCCACCACTGATCGCAACTCCATTTGAGGAAGGCGAGTA 254

Qy        216 TCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCT 275
              ||      || || || | ||||| ||     |||| ||||   ||||||||||  || |
Db        255 CCCAACGCTTCTACTCCTTCATGGATTCATGCTTCACAACACTTTCTACTCTGAGCTTAT 314

Qy        276 TAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCC 335
              | |||||||| || ||||| |  ||| |||| |  || ||||| |      ||   |
Db        315 CCAGCACATCGCATCCCATGGCTTTATTGTTGTTGTACCTCAGTTATATCTTGTAGCTAC 374

Qy        336 GCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAA 395
              |   |       |  |  | ||||| ||     | | | |||     |||
Db        375 ATGTGATAGTACGAATGGCATCAAGTCTGCTGCAAAAACAACAGATTGGTTGAAGGATGG 434

Qy        396 CCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGG 455
              ||  ||| |  || |||||   || |   |  |  ||| |  |  |||    ||
Db        435 ACTGCAAGATGTTCTCCCAACAAAAGTCAGACCAGACCTAAAGAAACTCGGACTGAGCGG 494

Qy        456 CCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGA 515
              || ||||| || || |||   || |||||  |||| |||||  ||||         |
Db        495 ACATAGCCGTGGCGGCAAAGATGCATTTGCTCTTGCACTAGGATATGC---------GAA 545

Qy        516 CCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATA 575
              |  |  | |  || ||||| || || || || ||||| || |  |||       ||
Db        546 GACTACATTAAGCTTCTCAGCGCTGATCGGGATCGATCCTGTTGACGGAGTACGAAAGGG 605

Qy        576 CATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACC 635
              ||||  |||| | | |  | |  ||||  ||  | ||||| || ||  | || ||
Db        606 ACACCAAACCAATCCTCCTGTACTGAATTATATCCCTCACTCTTTGGAACTCAAGATGCC 665

Qy        636 GGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGC 695
              |      |||  | ||||| ||  | ||        |  ||  ||| |  |     | || ||
Db        666 GTCATTAGTTATCGGAACAGGTTTAGGTGAATTGAAGAGGAACCTATT---TGCCTGTGC 722

Qy        696 ACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTT 755
              || | ||   |||| ||  | || |||||| | ||||||      |  ||      ||||||
Db        723 CCCTAAGGGAGTAAATCACCAAGATTTTTACGATGAGTGTTCTTCCCCGTCTTGCCATTT 782

Qy        756 CGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGG 815
              || |||  ||| || || || ||||| ||| | || ||| ||      |||  |   ||
Db        783 TGTTGCTAAGGACTATGGTCACATGGACATGCTAGATGATCATACTGACGGACCTAGAGG 842
```

```
Qy        816 GTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTT 875
              || || |    || ||||||   |||||  ||  || ||   |    ||| |    ||
Db        843 TATTGTGTCATACTGCATGTGTCGAAATGGAGAATCAAGAA---CACCCATGCGACTGTT 899

Qy        876 TGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGAT 935
              ||||||||| || |||||||| ||||| ||      | ||     |||| ||| |||  |
Db        900 TGTAGGTGGGATCGTGGTTGCTTTTCTGAAAGCCTGCTTACAAGGTGATAAACAGGATTT 959

Qy        936 TCGATTGATTGTGAAGGATCCTTCCGTTTC 965
              |||    |   | |  | ||    ||||
Db        960 GCGAGCTCTGAAGGATAACCCCGAGATTTC 989


RESULT 14
AAD38075
ID   AAD38075 standard; cDNA; 1216 BP.
XX
AC   AAD38075;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Grape chlorophyllase cDNA from vrl1c.pk008.o21:fis clone.
XX
KW   Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Vitis sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .938
FT                   /*tag=  a
FT                   /product= "Grape chlorophyllase protein from
FT                   vrl1c.pk008.o21:fis clone"
FT                   /EC_number= "3.1.1.14"
FT                   /note= "CDS does not include start codon"
FT                   /partial
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23777.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 48-49; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is grape
CC   chlorophyllase cDNA from vrl1c.pk008.o21:fis clone
XX
SQ   Sequence 1216 BP; 331 A; 275 C; 256 G; 354 T; 0 U; 0 Other;

  Query Match             14.3%;  Score 169.8;  DB 6;  Length 1216;
  Best Local Similarity   53.2%;  Pred. No. 2.1e-40;
  Matches  435;  Conservative    0;  Mismatches  367;  Indels   15;  Gaps    3;

Qy        167 CACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTT 226
              | || |||   |   |   ||  |   ||  | |  || ||   |  ||  |  |
Db        112 CCCCTCCACTGCAACTCTTGATTGCTACGCCCTCTGAAGCAGGGGAGTTCCCGCTGCTGC 171

Qy        227 TATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCG 286
              |  || | ||||| | |  |||     ||||  ||||||||  |  |  |  | ||||| |
Db        172 TCCTCCTTCATGGTTATCTTCTCTATAACTCTTTCTACTCCCAGCTCATCCAACACATAG 231

Qy        287 CTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGC 346
              | || ||||||| |||| || | || || ||||| | ||   |        || |
Db        232 CCTCTCATGGTTTCATTGTTCTTGCTCCTCAGTTATACACTGTGGCTGGACCAGATTCAA 291

Qy        347 AAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTC 406
              |  ||  |  |  || |   | | | ||| |||  |||  ||   ||| | |
Db        292 GCGAAGAGATCAAGTCCGCAGCTGCTTTAACAAATTGGTTATCCAAAGGACTCCATGACT 351

Qy        407 ACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCG 466
              || ||  |    ||     | |||  ||   |      ||| |  ||||| || || |
Db        352 TACTTCCTCCCCATGTTCGGCCAAATTTAAGCAAACTAGGACTTGCCGGCCATAGTCGTG 411

Qy        467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCA 526
              | || ||||| || |||||  | || ||  |          || |  |  | | ||  | |
Db        412 GAGGCAAAACTGCTTTTGCTCTAGCACTGAG---------AAAAGCATCCACTTCTCTGA 462

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
              ||||||||  | ||||| || || || || |  ||||   |||||    |    ||||
Db        463 AATTTTCAGCCTTGATAGGCATAGACCCGGTGGATGGAATGGACAAAGGGAAACAAACCC 522

Qy        587 ATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGG 646
              ||| |   |  | || |||   ||  | || || || || ||||||  | ||    ||
Db        523 CTCCACCGGTACTCACCTATGTTCCTCATTCATTTGATCTAGACATGGCAGTGATGGTAA 582

Qy        647 TGGGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGG 703
              | ||  | ||  | ||   |  |||  |||||    || | || || || || | || ||
Db        583 TTGGTTCGGGTTTGGGTGAAGTGAAAAGGAACCCTCTGTTCCCTCCTTGTGCCCCCAAGG 642

Qy        704 ACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTG 763
              | |||||||||||||| || |  ||||| ||     |   |         ||||  | ||
Db        643 GCGTAAACCATGAGGACTTCTTTAAAGAATGCCGTGAACCAGCTTGTTATTTTCTTGCCA 702

Qy        764 CGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGG 823
              ||| || || ||| | || ||| | ||||||||       ||  ||   |||
Db        703 AGGACTATGGCCATCTTGACATGCTAGACGATGAGACTAATGGAATTAGAGGGAAAGCTA 762

Qy        824 CCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTG 883
              |   |||| ||||||| |||||| |    | | |  |    |||||||| |||| || |
Db        763 CACATTGTTTGTGTAAAAATGGGAAGTCTAGAGAACC---CATGAGGAGGTTTGTTGGAG 819

Qy        884 GAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGA 943
              | |||||| |||| ||| | ||    |  ||||    || || ||        |    | |
Db        820 GCATTGTGATTGCATTTATGAAAGCTTATTTGGAAGGCGATAACAGCAGTCTAATCTCCA 879

Qy        944 TTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGA 980
              ||    | |  | | |   | | || |   |||| |
Db        880 TTAGAGATGGGCATGCTACTGCACCAGTGGAGCTTCA 916


RESULT 15
AAH41137
ID   AAH41137 standard; DNA; 1135 BP.
XX
AC   AAH41137;
XX
DT   21-AUG-2001  (first entry)
XX
DE   Arabidopsis thaliana chlorophyllase coding sequence #2.
XX
KW   Chlorophyllase; transgenic plant; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   JP2001086990-A.
```

XX
PD 03-APR-2001.
XX
PF 20-SEP-1999; 99JP-00266181.
XX
PR 20-SEP-1999; 99JP-00266181.
XX
PA (KAGO) KAGOME KK.
XX
DR WPI; 2001-338421/36.
DR P-PSDB; AAB99103.
XX
PT DNA encoding chlorophyllase, useful for producing transgenic plants.
XX
PS Claim 5; Page 12-14; 21pp; Japanese.
XX
CC The present sequence is a chlorophyllase protein coding sequence. The
CC chlorophyllase protein coding sequence can be used for the transformation
CC of a plant
XX
SQ Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 U; 0 Other;

Query Match 13.8%; Score 163.4; DB 5; Length 1135;
Best Local Similarity 54.3%; Pred. No. 1.8e-38;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

```
Qy        157 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 216
              | |||| |  | |||||||| | | ||     |  |     ||     |  |  |||  ||||
Db        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Qy        217 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 276
              || || ||  |  || |||||||  |     |||   ||||| ||||||  |||  |  || |
Db        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Qy        277 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 336
              ||  || |||| |||||  | |||  ||  | || || |||||  |  |    |
Db        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330

Qy        337 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 396
              || |         |  ||  |  |     | |       | ||  | |||  ||| | |
Db        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390

Qy        397 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 456
              || ||        || ||| |    ||||   | ||   |      |     | |||   |||
Db        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450

Qy        457 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 516
              || |||||||||||| ||||| |||||||||||| ||  ||    |     |     |
Db        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507

Qy        517 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 576
              |     | | | | ||  |  | || || || ||||||||||  |||||    |||
Db        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Qy        577 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 636
              |    ||||  ||| |   | ||  | ||    ||  | || || || || |||| | ||
Db        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Qy        637 GTTGCA---GTGGTGGGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 690
              |  |   ||| | ||  | || || ||   | |     |||||    |  | |||||
Db        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Qy        691 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 750
              || |||||  | |   | || ||    |||||  |      || ||| | |   |  |
Db        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Qy        751 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 810
              |||||||| ||   |||||| || ||| |||| ||| | || ||||||      ||  ||
Db        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807

Qy        811 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 870
              |||  |  |  |||| |||||||||||||  |   | | |  ||||||
Db        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864
```

```
Qy        871 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 923
              || || || |||||| |||| ||  | ||| | |||  |  ||||   || ||
Db        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
```

Search completed: March 20, 2006, 14:14:20
Job time : 704.094 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:06:13 ; Search time 1101.06 Seconds
(without alignments)
8922.325 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1188	100.0	1188	7	US-10-634-548-18	Sequence 18, Appl
2	975	82.1	975	3	US-09-938-842A-158	Sequence 158, App
3	975	82.1	975	3	US-09-938-842A-158	Sequence 158, App
4	748	63.0	1125	7	US-10-425-114-26342	Sequence 26342, A
5	239.4	20.2	987	9	US-10-381-123-1	Sequence 1, Appli
6	202.6	17.1	1174	9	US-10-381-123-7	Sequence 7, Appli
7	199	16.8	1242	9	US-10-381-123-17	Sequence 17, Appl
8	194.4	16.4	1104	9	US-10-381-123-9	Sequence 9, Appli
9	188	15.8	1125	9	US-10-381-123-11	Sequence 11, Appl
10	185.2	15.6	1307	7	US-10-424-599-45715	Sequence 45715, A
11	178.8	15.1	1444	9	US-10-381-123-13	Sequence 13, Appl
12	169.8	14.3	1216	9	US-10-381-123-3	Sequence 3, Appli
13	163.4	13.8	1135	7	US-10-634-548-19	Sequence 19, Appl
14	135.6	11.4	1068	7	US-10-424-599-34940	Sequence 34940, A
15	110.8	9.3	1156	9	US-10-381-123-25	Sequence 25, Appl
16	109.8	9.2	635	7	US-10-424-599-131410	Sequence 131410,
17	109.6	9.2	796	7	US-10-424-599-86252	Sequence 86252, A
18	109.2	9.2	751	8	US-10-425-115-70835	Sequence 70835, A
19	109.2	9.2	1302	9	US-10-381-123-5	Sequence 5, Appli
20	109.2	9.2	2439	8	US-10-425-115-140359	Sequence 140359,
21	107.8	9.1	538	7	US-10-424-599-34941	Sequence 34941, A
22	106.6	9.0	1244	7	US-10-425-114-25386	Sequence 25386, A
23	106.6	9.0	1244	7	US-10-425-114-25387	Sequence 25387, A
24	98.6	8.3	916	9	US-10-381-123-15	Sequence 15, Appl
25	82.6	7.0	1236	8	US-10-425-115-166593	Sequence 166593,
26	82.6	7.0	1274	9	US-10-381-123-27	Sequence 27, Appl
27	77.8	6.5	1004	7	US-10-424-599-68342	Sequence 68342, A
28	56.6	4.8	408	7	US-10-437-963-80160	Sequence 80160, A
29	53	4.5	1155	7	US-10-437-963-94232	Sequence 94232, A
30	52.6	4.4	728	7	US-10-767-701-3821	Sequence 3821, Ap
31	47	4.0	985	3	US-09-938-842A-2858	Sequence 2858, Ap
32	47	4.0	985	3	US-09-938-842A-2858	Sequence 2858, Ap
33	45	3.8	1829	3	US-09-974-879-60	Sequence 60, Appl
34	45	3.8	1829	3	US-09-305-736-60	Sequence 60, Appl
35	45	3.8	1829	3	US-09-818-683-60	Sequence 60, Appl
36	45	3.8	1829	3	US-09-818-683-60	Sequence 60, Appl
37	45	3.8	1829	7	US-10-621-401-60	Sequence 60, Appl
c 38	41.6	3.5	631	3	US-09-814-353-5993	Sequence 5993, Ap
c 39	41.6	3.5	631	3	US-09-814-353-12272	Sequence 12272, A
40	40.8	3.4	13376	6	US-10-311-455-556	Sequence 556, App
41	40.4	3.4	524	7	US-10-767-701-3830	Sequence 3830, Ap
c 42	39.8	3.4	752	3	US-09-814-353-18656	Sequence 18656, A
c 43	39.6	3.3	542	4	US-09-925-065A-175209	Sequence 175209,
44	39.6	3.3	604	5	US-10-106-698-1620	Sequence 1620, Ap
45	39.6	3.3	1345	8	US-10-723-860-5698	Sequence 5698, Ap

ALIGNMENTS

RESULT 1
US-10-634-548-18
; Sequence 18, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1ris, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen Kindle
; APPLICANT: Hirshberg, Joseph
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentin, Henry E.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
; FILE REFERENCE: Ren-01-125
; CURRENT APPLICATION NUMBER: US/10/634,548
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: us 60/400,689
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Synechocystis PCC6803
US-10-634-548-18

Query Match 100.0%; Score 1188; DB 7; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATACATAAATCTTCAACACAACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATA 60

Qy         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTC 120

Qy        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCG 180
```

```
Qy        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGC 240

Qy        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTAC 300

Qy        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGAC 360

Qy        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCG 420

Qy        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCG 480

Qy        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTA 540

Qy        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTA 600

Qy        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTC 660

Qy        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720

Qy        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATG 780

Qy        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAG 840

Qy        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTT 900

Qy        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCC 960

Qy        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTC 1020

Qy       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATG 1080

Qy       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAA 1140

Qy       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAAAAAAAAA 1188
```

RESULT 2
US-09-938-842A-158
; Sequence 158, Application US/09938842A
; Patent No. US20020160378A1

```
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 158
;   LENGTH: 975
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

  Query Match             82.1%;  Score 975;  DB 3;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-285;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023
              |||||||||||||||
Db        961 GGTATCTTCGTCTAG 975


RESULT 3
US-09-938-842A-158
; Sequence 158, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 158
;   LENGTH: 975
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

  Query Match             82.1%;  Score 975;  DB 3;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-285;
  Matches  975;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         49 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCGGCGATAGAGGACAGTCCAACGTTTTCCTCTGTGGTAACTCCGGCGGCTTTTGAG 60

Qy        109 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAGGCAGCCTCCCGACAACCGAGATACCGGTGGATCCGGTGGAAAATGATTCAACAGCA 120

Qy        169 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTTTA 180

Qy        229 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCGCT 240

Qy        289 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGCAA 300

Qy        349 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTCAC 360

Qy        409 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCGGT 420

Qy        469 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCACG 480

Qy        529 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCGAT 540

Qy        589 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGGTG 600

Qy        649 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGGACTTA 660

Qy        709 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGAT 720

Qy        769 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGT 780

Qy        829 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATT 840

Qy        889 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTG 900

Qy        949 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCT 960

Qy       1009 GGTATCTTCGTCTAG 1023
              |||||||||||||||
Db        961 GGTATCTTCGTCTAG 975


RESULT 4
US-10-425-114-26342
; Sequence 26342, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26342
;   LENGTH: 1125
;   TYPE: DNA
;   ORGANISM: Brassica napus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB4156-010-D3_FLI
```

US-10-425-114-26342

Query Match 63.0%; Score 748; DB 7; Length 1125;
Best Local Similarity 82.0%; Pred. No. 4e-216;
Matches 873; Conservative 0; Mismatches 190; Indels 1; Gaps 1;

```
Qy         21 AACTCTTTAATTATCTAGTTTAATACAAATGGCGGCGATAGAGGACAGTCCAACGTTTTC 80
Db         11 AACACACTCTTTTTCCAATATTATACAAATGGCGGGGAAGGAGGACAGTGAGACGTTTTA 70

Qy         81 CTCTGTGGTAACTCCGGCGGCTTTTGAGATAGGCAGCCTCCCGACAACCGAGATACCGGT 140
Db         71 CTCGGCGGCAACTCCTTTGGCGTTTGAGTTAGGCAGCCTTCCAACAACCGTGATCCCCGC 130

Qy        141 GGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAAC 200
Db        131 AGACCCGTCGGCAACCGATTTGACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAAC 190

Qy        201 AGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTT 260
Db        191 CGTCGCCGGAACTTACCCCGTCGTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTT 250

Qy        261 CTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTT 320
Db        251 CTACTCTGATGTTATTAACCACGTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCT 310

Qy        321 GTGCAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAA 380
Db        311 GTGCAAGATTTTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAA 370

Qy        381 CTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATA 440
Db        371 CTGGACTTCGAAAAACCTCAAAGCTCACCTCCCAAGCTCAGTAAACGCTAATGGCAACTA 430

Qy        441 CACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCA 500
Db        431 CACCGCACTCGTGGGCCATAGCCGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCA 490

Qy        501 TGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGC 560
Db        491 CGCCGCAACACTAGACCCATCCATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGC 550

Qy        561 AGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTT 620
Db        551 AGGAATCAGCAAATGCATAAGAACCGATCCCGAAATCTTAACGTACAAACCGGAATCATT 610

Qy        621 CGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGT 680
Db        611 CGACCTGGACATGCCGGTTGCAGTGATCGGTACGGGTCTCGGACCGAAGAGTAACATGCT 670

Qy        681 GATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGC 740
Db        671 GATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAGGC 730

Qy        741 GACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTT 800
Db        731 TACCAAGGGGCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATAATTT 790

Qy        801 GCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTC 860
Db        791 GCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACAGTAAACGCAAAAAGAG 850

Qy        861 TGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGG 920
Db        851 TGAGATGAGAAGCTTTGTTGGTGGAATTGTGGTTGCGTTTCTAAAGTATAGTCT-TTGGG 909

Qy        921 TGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGA 980
Db        910 TGAAAAGTCAGAGATTCGACAGATTGTGAAGGATCCCTCTGTTTCTCCGGCGAGGCTTGA 969

Qy        981 TCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTA 1040
Db        970 TCCTTTGCCGGAGCTGGAAGAGGCTTCTGGTTATCTCGTCTAGATTTGTGGTTAAGGAAT 1029
```

```
Qy       1041 TTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTT 1084
Db       1030 GTATTGGTATCCTCAGGGAAGCTCCAAACTCGTAATAAACTTAT 1073
```

RESULT 5
US-10-381-123-1
; Sequence 1, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Vitis sp.
US-10-381-123-1

Query Match 20.2%; Score 239.4; DB 9; Length 987;
Best Local Similarity 57.6%; Pred. No. 2e-61;
Matches 473; Conservative 0; Mismatches 336; Indels 12; Gaps 2;

```
Qy        167 CACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTT 226
Db         26 CCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTCTCT 85

Qy        227 TATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCG 286
Db         86 TGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCATTT 145

Qy        287 CTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGC 346
Db        146 CTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTGGAA 205

Qy        347 AAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTC 406
Db        206 TTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAATCTG 265

Qy        407 ACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCG 466
Db        266 TGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCAGAG 325

Qy        467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCA 526
Db        326 GGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGC---------TGATACATCCCTCA 376

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        587 ATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGG 646
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAGCAATCCCAGTTTGCGTAA 496

Qy        647 TGGGAACCGGACTCGGACCGAAG---TGGAACAACGTGATGCCACCATGCGCACCAACGG 703
Db        497 TCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAGATG 556

Qy        704 ACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTG 763
Db        557 GAGTGAACCATGTAGAGTTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGACTA 616

Qy        764 CGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGG 823
```

```
Db        617 CTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGATTT 676

Qy        824 CCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTG 883

Db        677 CGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGGGTG 736

Qy        884 GAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGA 943

Db        737 GCCTTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATTTCAAAGCCA 796

Qy        944 TTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCT 984

Db        797 TTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCT 837


RESULT 6
US-10-381-123-7
; Sequence 7, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 1174
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-381-123-7

  Query Match             17.1%;  Score 202.6;  DB 9;  Length 1174;
  Best Local Similarity   56.9%;  Pred. No. 3.5e-50;
  Matches  480;  Conservative    0;  Mismatches  339;  Indels   24;  Gaps    5;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213

Db        102 AATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTGGCTCA 161

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273

Db        162 TACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTGAGCTC 221

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTG 333

Db        222 CTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGAGTGTA 281

Qy        334 CCGCC------GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGG--- 384

Db        282 CGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATTGGCTA 341

Qy        385 GCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACC 444

Db        342 GCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATAAATTG 401

Qy        445 TCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCC 504

Db        402 GTTTTATCAGGTCACAGCAAGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTTATGCT 461

Qy        505 GCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGA 564

Db        462 AAAA---------CTAACCTCAAGTTTTCAGCACTAGTAGGCATAGACCCTGTGGCTGGC 512

Qy        565 ACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAG 624

Db        513 CCATGTAAATCTTGCGAAACATTTCCTCCTATTCTCACTGGCATGTCCCAATCCTTCAAT 572

Qy        625 CTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTG 681

Db        573 TTGAACATACCCATTGTTGTAATTGGCACTGGGCTAGGCCCAGAGAAGGCTAATTTTTTT 632

Qy        682 ATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCG 741

Db        633 ATTCCACCATGTGCTCCTGATGGGGTGAACCATAAGGAGTTTTTCAATAAGTGCAAACCC 692

Qy        742 ACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTG 801

Db        693 CCTTGTGCACATTTTGTTGCAACTGAGTATGGTCACATGGACATGTTGGATGATGTGACA 752

Qy        802 CCCGGTTTTGTTGGGT---TTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAG 858

Db        753 CCTGGCTTAATTGGGTCAATATTGTCAAATTGTATATGCAAGGATGGGAAGGGTCCTAGG 812

Qy        859 TCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGG 918

Db        813 GACTTGATGAGAAGGACCGTGGGAGGGTTGGTTGTGGCCTTCTTAAGGGCACAGTTGAAT 872

Qy        919 GGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTT 978

Db        873 GGCCTATGGAAGGATTTTAATGCTGTTTTGGCGAATCCTAATCTTGCTCCTACTAAACTG 932

Qy        979 GAT 981

Db        933 GAT 935


RESULT 7
US-10-381-123-17
; Sequence 17, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
;   LENGTH: 1242
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
US-10-381-123-17

  Query Match             16.8%;  Score 199;  DB 9;  Length 1242;
  Best Local Similarity   55.8%;  Pred. No. 4.5e-49;
  Matches  476;  Conservative    0;  Mismatches  350;  Indels   27;  Gaps    4;

Qy        137 CGGTGGATCCGGTGGAAAATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTC 196

Db        153 CGATTCAGGTGGATGAGAATGCGGCGCCGACACCACCGATCCCGGTGCTGATCGTCGCAC 212

Qy        197 CAACAGTCGCCGGAACTTATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACT 256

Db        213 CCAAGGATGCAGGAACCTACCCCGTGGCCATGCTCTTGCACGGCTTCTTCCTCCATAACC 272

Qy        257 ACTTCTACTCTGACGTTCTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCAC 316

Db        273 ACTTCTACGAACACCTTCTCCGGCACGTCGCATCCCACGGCTTCATCATTGTCGCGCCCC 332

Qy        317 AGTTGTG---CAAATTATTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTG 373

Db        333 AGTTCAGCATCAGTATCATACCTTCGGGTGACGCAGAGGACATCGCCGCGGCAGCCAAGG 392

Qy        374 TGATAAACTGGGCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATG 433
```

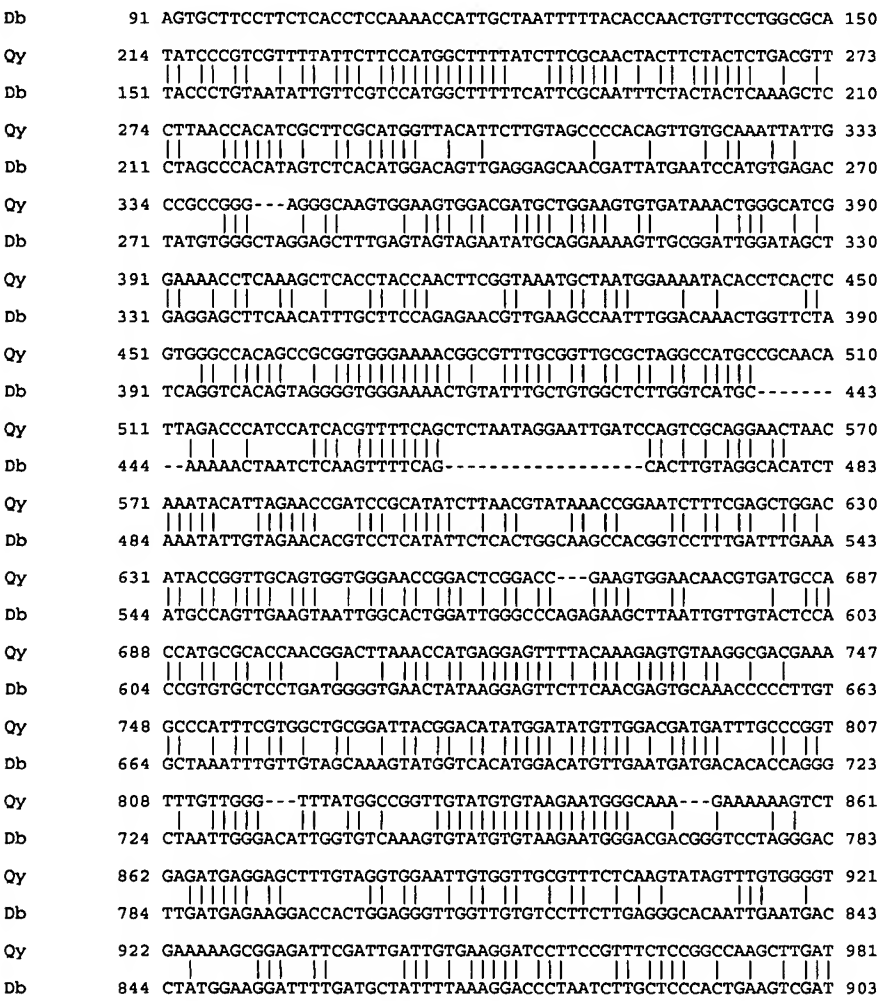
```
Db        393 TGGCAGACTGGCTCCCCGACGGCCTCCCGTCCGTGCTGCCCAAAGGCGTCGAGCCGGAGC 452

Qy        434 GAAAATACACCTCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGC 493

Db        453 TCTCGAAGCTCGCCTTGGCCGGCCACAGCCGAGGAGGCCACACGGCTTTCTCCCTGGCCT 512

Qy        494 TAGGCCATGCCGCAACATTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATC 553

Db        513 TGGGGCACGCC---------AAGACCCAGCTAACCTTCTCCGCGCTCATCGGACTCGACC 563

Qy        554 CAGTCGCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGG 613

Db        564 CCGTCGCCGGCACGGGGAAGTCCTCCCAGCTCCAGCCCAAGATCCTCACCTACGAGCCGT 623

Qy        614 AATCTTTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGA 673

Db        624 CCTCCTTCGGCATGGCGATGCCGGTGCTGGTCATCGGCACCGGGCTCGGCGAGGAGAAGA 683

Qy        674 ACAACGTGATG---CCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAG 730

Db        684 AGAACATATTCTTCCCTCCCTGCGCACCCAAGGACGTGAACCACGCCGAGTTCTACCGCG 743

Qy        731 AGTGTAAGGCGACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGG 790

Db        744 AGTGCAGGCCGCCCTGCTACTACTTTGTGACCAAGGACTACGGGCATCTGGACATGCTGG 803

Qy        791 ACGATGATTTGCCCGGTTTTGTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA 850

Db        804 ACGACGACGCCCCCAAGTTCAT------------CACCTGCGTCTGCAAGGATGGGAACG 851

Qy        851 GAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATA 910

Db        852 GGTGCAAGGGCAAGATGCGGAGGTGCGTTGCTGGGATCATGGTGGCATTTCTTAATGCTG 911

Qy        911 GTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGG 970

Db        912 CCTTGGGTGAGAAAGATGCAGATCTTGAGGCCATACTGAGAGACCCGGCGGTTGCACCCA 971

Qy        971 CCAAGCTTGATCC 983

Db        972 CCACGCTTGATCC 984


RESULT 8
US-10-381-123-9
; Sequence 9, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
;   LENGTH: 1104
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-381-123-9

  Query Match            16.4%;  Score 194.4;  DB 9;  Length 1104;
  Best Local Similarity   56.3%;  Pred. No. 1e-47;
  Matches  473;  Conservative    0;  Mismatches  346;  Indels   21;  Gaps    5;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213

Db        113 AGTGCTTCCTTCTCACCTCCAAAACCATTGCTAATTTTTACACCAACTGTTCCTGGCGCA 172

Qy        214 TATCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273

Db        173 TACCCTGTAATATTGTTCGTCCATGGCTTTTTCATTCGCAATTTCTACTACTCAAAGCTC 232

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCC---ACAGTTGTGCAAATTA 330

Db        233 CTAGCCCACATAGTCTCACATGGATTCATAATCGTTGCTCCTCAACTGTTTTCCAATGGG 292

Qy        331 TTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCG 390

Db        293 CTTCCTATGTATGGACCCACTGAAGTGGAATATGCAGGAAAAGTTGCGGATTGGATAGCT 352

Qy        391 GAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTC 450

Db        353 GAGGAGCTTCAACATTTGCTTCCAGAGAACGTTGAAGCCAATTTGGACAAACTGGTTCTA 412

Qy        451 GTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACA 510

Db        413 TCAGGTCACAGTAGGGGTGGGAAAACTGTATTTGCTGTGGCTCTTGGTCATGC------- 465

Qy        511 TTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAAC 570

Db        466 --AAAAACTAATCTCAAGTTTTCAGCACTTGTAGGCATAGACCCTGTGGCTGGCACATCT 523

Qy        571 AAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGAC 630

Db        524 AAATATTGTAGAACACGTCCTCATATTCTCACTGGCAAGCCACGGTCCTTTGATTTGAAA 583

Qy        631 ATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTGATGCCA 687

Db        584 ATGCCAGTTGAAGTAATTGGCACTGGATTGGGCCCAGAGAAGCTTAATTGTTGTACTCCA 643

Qy        688 CCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAA 747

Db        644 CCGTGTGCTCCTGATGGGGTGAACTATAAGGAGTTCTTCAACGAGTGCAAACCCCCTTGT 703

Qy        748 GCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGT 807

Db        704 GCTAAATTTGTTGTAGCAAAGTATGGTCACATGGACATGTTGAATGATGACACACCAGGG 763

Qy        808 TTTGTTGGG---TTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA---GAAAAAAGTCT 861

Db        764 CTAATTGGGACATTGGTGTCAAAGTGTATGTGTAAGAATGGGACGACGGGTCCTAGGGAC 823

Qy        862 GAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGT 921

Db        824 TTGATGAGAAGGACCACTGGAGGGTTGGTTGTGTCCTTCTTGAGGGCACAATTGAATGAC 883

Qy        922 GAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGAT 981

Db        884 CTATGGAAGGATTTTGATGCTATTTTAAAGGACCCTAATCTTGCTCCCACTGAAGTCGAT 943


RESULT 9
US-10-381-123-11
; Sequence 11, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
;   LENGTH: 1125
;   TYPE: DNA
```

```
;  ORGANISM: Glycine max
US-10-381-123-11

  Query Match             15.8%;  Score 188;  DB 9;  Length 1125;
  Best Local Similarity   56.6%;  Pred. No. 9.4e-46;
  Matches  479;  Conservative    0;  Mismatches  340;  Indels   27;  Gaps    6;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
Db        132 ACTGCTTCTTCCTCACCTCCAAAACCATTGCTTATTTTACACCAACCGTGCCTGGCTTA 191

Qy        214 TATCCCGTCGTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
Db        192 TACCCTGTAATATTGTTTGCCATGGCTTTTGTATTCGCACTAGCTACTACTCTAAGCTC 251

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACA---GTTGTGCAAATTA 330
Db        252 CTAGCCCACATAGTTTCACATGGATTCATACTTGTTGCTCCTCAGCTGTTTTCCATTGGG 311

Qy        331 TTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCG 390
Db        312 GTGCCTATGTTTGGACCAGAAGAAGTTAAGTGTGAAGGAAGAGTTGTGGATTGGCTAGAT 371

Qy        391 GAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTC 450
Db        372 AACGGGCTTCAACCATTGCTTCCCGAGAGCGTTGAAGCCAAACTGGAGAAACTGGTTCTA 431

Qy        451 GTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACA 510
Db        432 GTAGGTCACAGCAAGGGTGGAAAAACAGCATTTGCTGTGGCACTTGGTTA---------C 482

Qy        511 TTAGACCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGG------A 564
Db        483 TGTAAAACAAAGCTCAAGTTTTCAGCACTCATAGGCATAGATCCTGTGGCTGGCGTATCA 542

Qy        565 ACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAG 624
Db        543 AAGTGTAAGCCTTGTCGATCACTTCCTGATATCCTCACAGGTGTGCCACGGTCCTTTAAT 602

Qy        625 CTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGTG 681
Db        603 CTGAACATACCTGTTGCTGTAATTGGAACTGGGTTGGGCCCAGAGAAGGCTAATTCTCTT 662

Qy        682 ATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCG 741
Db        663 TTTCCACCATGTGCTCCAAATGGGGTGAACCATAAAGAGTTTTTCTCTGAGTGCAAACCA 722

Qy        742 ACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTG 801
Db        723 CCTAGTGCATATTTTGTTGCAACGGATTATGGTCACATGGACATGTTGGATGATGAAACA 782

Qy        802 CCCGGTTTTGTTGGG---TTTATGGCCGGTTGTATGTGTAAGAATGGGCAAA---GAAAA 855
Db        783 CCAGGGGTAATTGGGACAATGATGTCAAAGTGTATGTGCAAGAATGGGAAGAAGGGTCCT 842

Qy        856 AAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTG 915
Db        843 AGGGACTTGATGAGAAGGACTGTGGGAGGGTTGGTTGTGGCCTTCTTGAGAGCACAGTTG 902

Qy        916 TGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAG 975
Db        903 AATGAGCAGTGGAAGGATTTTGATGCTATTTTAGCGAGTCCTAATCTAGCTCCCGCCAAA 962

Qy        976 CTTGAT 981
Db        963 CTTGAT 968

RESULT 10
US-10-424-599-45715
; Sequence 45715, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 45715
;   LENGTH: 1307
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_141281C.1
US-10-424-599-45715

  Query Match             15.6%;  Score 185.2;  DB 7;  Length 1307;
  Best Local Similarity   56.4%;  Pred. No. 7.3e-45;
  Matches  476;  Conservative    0;  Mismatches  343;  Indels   25;  Gaps    6;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
Db        156 AATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTGGCTCA 215

Qy        214 TATCCCGTCGTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTT 273
Db        216 TACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTGAGCTC 275

Qy        274 CTTAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTG 333
Db        276 CTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGAGTGTA 335

Qy        334 CCGCC------GGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGG--- 384
Db        336 CGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATTGGCTA 395

Qy        385 GCATCGGAAAACCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACC 444
Db        396 GCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATAAATTG 455

Qy        445 TCACTCGTGGGCCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCC 504
Db        456 GTTTTATCAGGTCACACCTTGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTTATGCT 515

Qy        505 GCAACATTAGACCCATCCATCACGTTTTCAGCTCTAAT-AGGAATTGATCCAGTCGCAGG 563
Db        516 AAAA---------CTGACCTCAAGTTGTCAGCACTAGTCAGGCATAGACCCTGTGGCTGG 566

Qy        564 AACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGA 623
Db        567 CCCATGTAAATCTTGCGAAACATTTCCTCCTATTCTCACTGGCATGTCCCAATCCTTCAA 626

Qy        624 GCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACC---GAAGTGGAACAACGT 680
Db        627 TTTGAACATACCCATTGTTGTAATTGGCACTGGGCTAGGCCCAGAGAAGGCTAATTTTTT 686

Qy        681 GATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGC 740
Db        687 TATTCCACCATGTGCTCCTGATGGGGTGAACCATAAGGAGTTTTTCAATAAGTGCAAACC 746

Qy        741 GACGAAAGCCCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTT 800
Db        747 CCCTTGTGCACATTTTGTTGCAACTGAGTATGGTCACATGGACATGTTGGATGATGTGAC 806

Qy        801 GCCCGGTTTTGTTGGGT---TTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAA 857
Db        807 ACCTGGCTTAATTGGGTCAATATTGTCAAATTGTATATGCAAGGCCGGGAAGGGTCCTAG 866

Qy        858 GTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTG 917
Db        867 GGACTTGATGAGAAGGACCGTGGGAGGGTTGGTTGTGGCCTTCTTAAGGGCACAGTTGAA 926
```

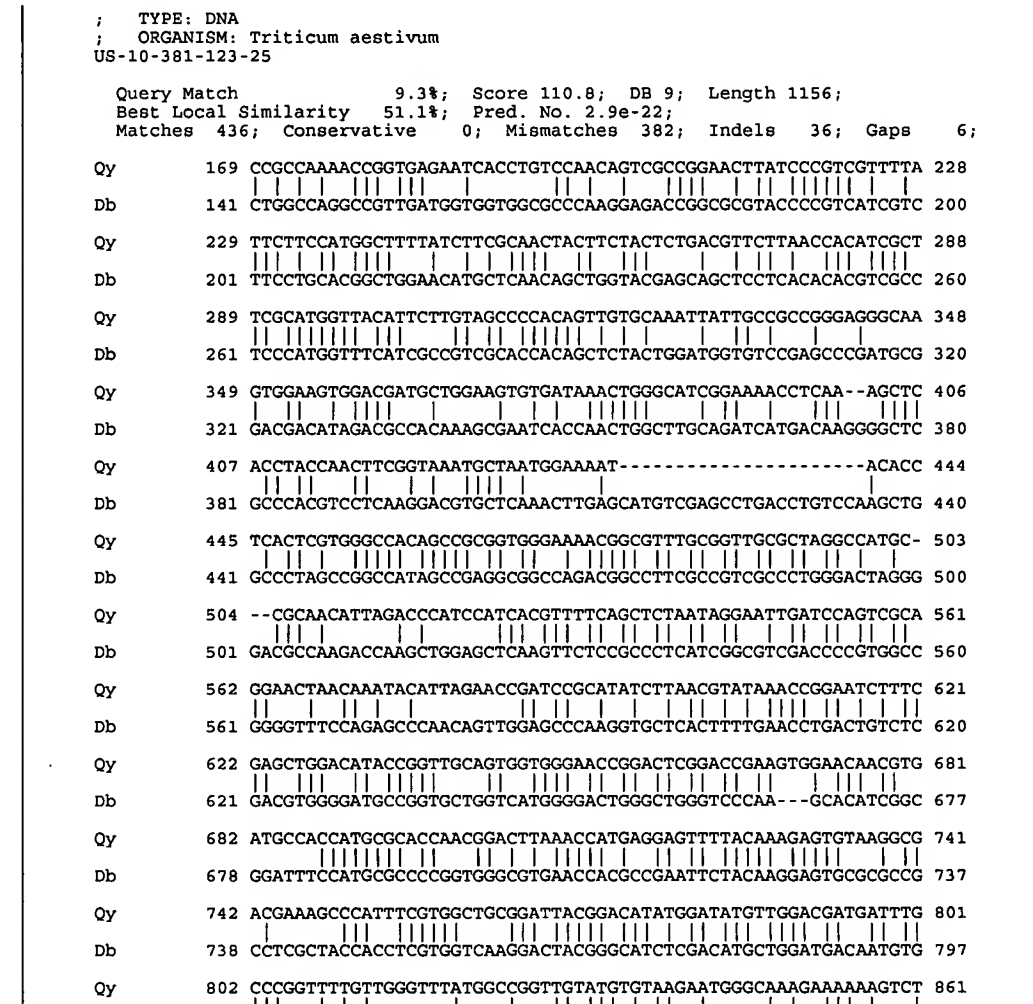
```
Qy        918 GGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCT 977
              ||    |    |||  ||      || ||  | |||||    || ||||  | || ||
Db        927 TGGCCTATGGAAGGATTTTAATGCTGTTTTGGCGAATCCTAATCTTGCTCCTACTAAACT 986

Qy        978 TGAT 981
              |||
Db        987 GGAT 990


RESULT 11
US-10-381-123-13
; Sequence 13, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
;   LENGTH: 1444
;   TYPE: DNA
;   ORGANISM: Gesneriana
US-10-381-123-13

  Query Match             15.1%;  Score 178.8;  DB 9;  Length 1444;
  Best Local Similarity   54.1%;  Pred. No. 6.8e-43;
  Matches  438;  Conservative    0;  Mismatches  357;  Indels   15;  Gaps    3;

Qy        156 TGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTA 215
              ||  ||  || |||| |  || ||       ||| |   ||||   |  |  ||    ||
Db        195 TGTCTCCGCACCACCTCGGAAGCCACCACTGATCGCAACTCCATTTGAGGAAGGCGAGTA 254

Qy        216 TCCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCT 275
              ||     || || || | ||||| ||    |||| ||||   ||||||||||  || |
Db        255 CCCAACGCTTCTACTCCTTCATGGATTCATGCTTCACAACACTTTCTACTCTGAGCTTAT 314

Qy        276 TAACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCC 335
              | |||||||| || ||||| |  ||| |||| |  || ||||| |       ||   |
Db        315 CCAGCACATCGCATCCCATGGCTTTATTGTTGTTGTACCTCAGTTATATCTTGTAGCTAC 374

Qy        336 GCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAA 395
              |   |        |  |  |  ||||| ||     | | | |||     |||
Db        375 ATGTGATAGTACGAATGGCATCAAGTCTGCTGCAAAAACAACAGATTGGTTGAAGGATGG 434

Qy        396 CCTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGG 455
              ||  ||| |   || |||||    || |   | |  ||| |   |  |||    ||
Db        435 ACTGCAAGATGTTCTCCCAACAAAAGTCAGACCAGACCTAAAGAAACTCGGACTGAGCGG 494

Qy        456 CCACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGA 515
              || ||||| || || |||   || |||||  |||| |||||  ||||              |
Db        495 ACATAGCCGTGGCGGCAAAGATGCATTTGCTCTTGCACTAGGATATGC---------GAA 545

Qy        516 CCCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATA 575
              |  |  | |  || ||||| || || || || ||||| || |  |||      ||
Db        546 GACTACATTAAGCTTCTCAGCGCTGATCGGGATCGATCCTGTTGACGGAGTACGAAAGGG 605

Qy        576 CATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACC 635
              |||| |||| | | |  | |  |||| ||  | ||||| || ||  | || ||
Db        606 ACACCAAACCAATCCTCCTGTACTGAATTATATCCCTCACTCTTTGGAACTCAAGATGCC 665

Qy        636 GGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCATGCGC 695
              |   |||  | ||||| ||  | ||     |  ||  ||| |  |    | || ||
Db        666 GTCATTAGTTATCGGAACAGGTTTAGGTGAATTGAAGAGGAACCTATT---TGCCTGTGC 722

Qy        696 ACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTT 755
```

```
              || | ||   |||| ||  | || |||||| | ||||||     |  ||     ||||||
Db        723 CCCTAAGGGAGTAAATCACCAAGATTTTTACGATGAGTGTTCTTCCCCGTCTTGCCATTT 782

Qy        756 CGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGG 815
              || |||   ||| || || || ||||| ||| | || ||| ||      |||    |  ||
Db        783 TGTTGCTAAGGACTATGGTCACATGGACATGCTAGATGATCATACTGACGGACCTAGAGG 842

Qy        816 GTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTT 875
              || || |    || ||||||   |||||  ||  || ||   |    ||| |    ||
Db        843 TATTGTGTCATACTGCATGTGTCGAAATGGAGAATCAAGAA---CACCCATGCGACTGTT 899

Qy        876 TGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGAT 935
              ||||||||| || |||||||| ||||| ||      | ||    |||||  |||  ||| |
Db        900 TGTAGGTGGGATCGTGGTTGCTTTTCTGAAAGCCTGCTTACAAGGTGATAAACAGGATTT 959

Qy        936 TCGATTGATTGTGAAGGATCCTTCCGTTTC 965
              |||     |   | |  | ||     ||||
Db        960 GCGAGCTCTGAAGGATAACCCCGAGATTTC 989


RESULT 12
US-10-381-123-3
; Sequence 3, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
;   LENGTH: 1216
;   TYPE: DNA
;   ORGANISM: Vitis sp.
US-10-381-123-3

  Query Match             14.3%;  Score 169.8;  DB 9;  Length 1216;
  Best Local Similarity   53.2%;  Pred. No. 3.4e-40;
  Matches  435;  Conservative    0;  Mismatches  367;  Indels   15;  Gaps    3;

Qy        167 CACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCCCGTCGTTT 226
              | || |||   |  |    ||  |    ||  | |  || ||    |  ||  |  |
Db        112 CCCCTCCACTGCAACTCTTGATTGCTACGCCCTCTGAAGCAGGGGAGTTCCCGCTGCTGC 171

Qy        227 TATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAACCACATCG 286
              |  || | ||||| | |  |||    ||||  ||||||||  |  |  |  | ||||| |
Db        172 TCCTCCTTCATGGTTATCTTCTCTATAACTCTTTCTACTCCCAGCTCATCCAACACATAG 231

Qy        287 CTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCGCCGGGAGGGC 346
              | || ||||||| |||| || | || || ||||| | ||   |       || |
Db        232 CCTCTCATGGTTTCATTGTTCTTGCTCCTCAGTTATACACTGTGGCTGGACCAGATTCAA 291

Qy        347 AAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAACCTCAAAGCTC 406
              |  ||  |  |    || |    | | | ||| |||  |||  ||   ||| | |
Db        292 GCGAAGAGATCAAGTCCGCAGCTGCTTTAACAAATTGGTTATCCAAAGGACTCCATGACT 351

Qy        407 ACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGCCACAGCCGCG 466
              || ||  |    ||    | |||  ||  |     ||| |  ||||| || || |
Db        352 TACTTCCTCCCCATGTTCGGCCAAATTTAAGCAAACTAGGACTTGCCGGCCATAGTCGTG 411

Qy        467 GTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGACCCATCCATCA 526
              | || ||||| || ||||   | || ||  |          || |  |  | | || | |
Db        412 GAGGCAAAACTGCTTTTGCTCTAGCACTGAG---------AAAAGCATCCACTTCTCTGA 462

Qy        527 CGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATACATTAGAACCG 586
              ||||||||  | ||||| || || || || | ||||   |||||    |    ||||
Db        463 AATTTTCAGCCTTGATAGGCATAGACCCGGTGGATGGAATGGACAAAGGGAAACAAACCC 522
```

```
Qy        587 ATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCGGTTGCAGTGG 646
Db        523 CTCCACCGGTACTCACCTATGTTCCTCATTCATTTGATCTAGACATGGCAGTGATGGTAA 582

Qy        647 TGGGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCATGCGCACCAACGG 703
Db        583 TTGGTTCGGGTTTGGGTGAAGTGAAAAGGAACCCTCTGTTCCCTCCTTGTGCCCCCAAGG 642

Qy        704 ACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCCATTTCGTGGCTG 763
Db        643 GCGTAAACCATGAGGACTTCTTTAAAGAATGCCGTGAACCAGCTTGTTATTTTCTTGCCA 702

Qy        764 CGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTGTTGGGTTTATGG 823
Db        703 AGGACTATGGCCATCTTGACATGCTAGACGATGAGACTAATGGAATTAGAGGGAAAGCTA 762

Qy        824 CCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTG 883
Db        763 CACATTGTTTGTGTAAAAATGGGAAGTCTAGAGAACC---CATGAGGAGGTTTGTTGGAG 819

Qy        884 GAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGA 943
Db        820 GCATTGTGATTGCATTTATGAAAGCTTATTTGGAAGGCGATAACAGCAGTCTAATCTCCA 879

Qy        944 TTGTGAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGA 980
Db        880 TTAGAGATGGGCATGCTACTGCACCAGTGGAGCTTCA 916


RESULT 13
US-10-634-548-19
; Sequence 19, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
;  APPLICANT: No. US20040045051A1ris, Susan R
;  APPLICANT:  Lincoln, Kim
;  APPLICANT:  Abad, Mark Scott
;  APPLICANT:  Eilers, Robert
;  APPLICANT:  Hartsuyker, Karen Kindle
;  APPLICANT:  Hirshberg, Joseph
;  APPLICANT:  Karunanandaa, Balasulojini
;  APPLICANT:  Moshiri, Farhad
;  APPLICANT:  Stein, Joshua C.
;  APPLICANT:  Valentin, Henry E.
;  APPLICANT:  Venkatesh, Tyamagondlu V.
;  TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
;  FILE REFERENCE: Ren-01-125
;  CURRENT APPLICATION NUMBER: US/10/634,548
;  CURRENT FILING DATE:  2003-08-05
;  PRIOR APPLICATION NUMBER: us 60/400,689
;  PRIOR FILING DATE: 2002-08-05
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
;   LENGTH: 1135
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-634-548-19

  Query Match             13.8%;  Score 163.4;  DB 7;  Length 1135;
  Best Local Similarity   54.3%;  Pred. No. 2.9e-38;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;

Qy        157 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 216
Db        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Qy        217 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 276
Db        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Qy        277 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 336
Db        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330

Qy        337 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 396
Db        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390

Qy        397 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 456
Db        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450

Qy        457 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 516
Db        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507

Qy        517 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 576
Db        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Qy        577 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 636
Db        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Qy        637 GTTGCA---GTGGTGGGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 690
Db        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Qy        691 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 750
Db        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Qy        751 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 810
Db        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807

Qy        811 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 870
Db        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864

Qy        871 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 923
Db        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917


RESULT 14
US-10-424-599-34940
; Sequence 34940, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 34940
;   LENGTH: 1068
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_131553C.1
US-10-424-599-34940

  Query Match             11.4%;  Score 135.6;  DB 7;  Length 1068;
  Best Local Similarity   53.2%;  Pred. No. 7.9e-30;
  Matches  447;  Conservative    0;  Mismatches  354;  Indels   39;  Gaps    6;

Qy        154 AATGATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACT 213
```



RESULT 15
US-10-381-123-25
; Sequence 25, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 1156



Search completed: March 20, 2006, 14:42:17
Job time : 1105.06 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:07:47 ; Search time 475.098 Seconds
(without alignments)
5830.519 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	39.6	3.3	542	6	US-09-925-065A-175209	Sequence 175209,
c	2	38.6	3.2	958	6	US-09-925-065A-683561	Sequence 683561,
	3	37.6	3.2	534	6	US-09-925-065A-102394	Sequence 102394,
	4	37.6	3.2	543	6	US-09-925-065A-849244	Sequence 849244,
c	5	37.4	3.1	605	6	US-09-925-065A-902549	Sequence 902549,
	6	37.2	3.1	578	6	US-09-925-065A-358029	Sequence 358029,
	7	37	3.1	590	6	US-09-925-065A-106455	Sequence 106455,
c	8	37	3.1	594	6	US-09-925-065A-417521	Sequence 417521,
	9	37	3.1	599	6	US-09-925-065A-106454	Sequence 106454,
c	10	37	3.1	893	6	US-09-925-065A-283413	Sequence 283413,
	11	37	3.1	5666	8	US-10-240-708-29	Sequence 29, Appl
	12	36.4	3.1	1691140	12	US-11-091-018-1	Sequence 1, Appli
	13	36	3.0	617	6	US-09-925-065A-881350	Sequence 881350,
c	14	36	3.0	1462	8	US-10-750-185-46341	Sequence 46341, A
c	15	36	3.0	1462	8	US-10-750-623-46341	Sequence 46341, A
	16	36	3.0	1830	8	US-10-750-185-38673	Sequence 38673, A
	17	36	3.0	1830	8	US-10-750-623-38673	Sequence 38673, A
c	18	35.6	3.0	179487	7	US-10-330-773-664	Sequence 664, App
	19	35.2	3.0	614	6	US-09-925-065A-308174	Sequence 308174,
c	20	35.2	3.0	616	6	US-09-925-065A-299192	Sequence 299192,
	21	35.2	3.0	1144	9	US-11-096-568A-1623	Sequence 1623, Ap
c	22	35.2	3.0	3769	8	US-10-793-626-4106	Sequence 4106, Ap
c	23	35.2	3.0	4249	8	US-10-793-626-4133	Sequence 4133, Ap
	24	35	2.9	201	8	US-10-995-561-22763	Sequence 22763, A
	25	35	2.9	1085	8	US-10-775-169-78	Sequence 78, Appl
	26	35	2.9	1233	12	US-11-098-686-9335	Sequence 9335, Ap
	27	35	2.9	3581	8	US-10-750-185-52094	Sequence 52094, A
	28	35	2.9	3581	8	US-10-750-623-52094	Sequence 52094, A
c	29	35	2.9	4356	7	US-10-932-182A-2957	Sequence 2957, Ap
c	30	35	2.9	4356	7	US-10-932-182A-2957	Sequence 2957, Ap
c	31	35	2.9	10144	8	US-10-240-708-93	Sequence 93, Appl
	32	35	2.9	317876	8	US-10-995-561-13227	Sequence 13227, A
	33	35	2.9	1457619	12	US-11-098-686-8739	Sequence 8739, Ap
c	34	34.8	2.9	550	6	US-09-925-065A-256291	Sequence 256291,
	35	34.6	2.9	464	6	US-09-925-065A-383958	Sequence 383958,
c	36	34.6	2.9	567	6	US-09-925-065A-655069	Sequence 655069,
	37	34.6	2.9	1400	12	US-11-136-527-4230	Sequence 4230, Ap
	38	34.6	2.9	2273	12	US-11-136-527-134	Sequence 134, App
c	39	34.6	2.9	194186	7	US-10-330-773-395	Sequence 395, App
	40	34.6	2.9	225587	7	US-10-330-773-374	Sequence 374, App
	41	34.4	2.9	559	6	US-09-925-065A-257742	Sequence 257742,
	42	34.4	2.9	623	6	US-09-925-065A-279572	Sequence 279572,
c	43	34.4	2.9	1496	8	US-10-750-185-63067	Sequence 63067, A
c	44	34.4	2.9	1496	8	US-10-750-623-63067	Sequence 63067, A
	45	34.2	2.9	201	8	US-10-995-561-51601	Sequence 51601, A

ALIGNMENTS

RESULT 1
US-09-925-065A-175209/c
; Sequence 175209, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 175209
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-175209

Query Match 3.3%; Score 39.6; DB 6; Length 542;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```
Qy       1055 TTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCC 1114
             || |  |||||| ||| |     | ||      |      ||  ||||| ||  |  | |
Db        384 TTTATGATTTGAGAAAATTTCATTTTTTAAAGATAATGAATTTAGCTCTTCCGCTAGGGC 325

Qy       1115 TAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAA 1174
              | |||   | | |||| |    |||| | || |||   ||| || |  ||||      |
Db        324 CATGTTCAGTCTTTATTCTGGCAAAACACAATGAAAGGTTTTTTTTTTTTTTTTTTTTA 265

Qy       1175 AAAAAAAAAAAAAA 1188
              ||||||||||||||
```

```
Db        264 AAAAAAAAAAAAAA 251


RESULT 2
US-09-925-065A-683561/c
; Sequence 683561, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 683561
;   LENGTH: 958
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-683561

  Query Match             3.2%;  Score 38.6;  DB 6;  Length 958;
  Best Local Similarity   47.7%;  Pred. No. 2.6;
  Matches  113;  Conservative    0;  Mismatches  124;  Indels    0;  Gaps    0;

Qy        854 AAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTT 913
              |||    |  |  |||  |   || || ||| ||      || | |         ||| |
Db        938 AAAGCATTATGCCGAGATGTCATGAAGATGGGATGTGTAGTGAGCTATCAGTAGCTAGAT 879

Qy        914 TGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCGGCCA 973
                | ||  | ||    |  || |    ||||| |||| | | | |  |||||| |
Db        878 GCTTGGTGGGAATTTGGTTGACTTACATGATTATGAATTAGCATCCTATTTCTCTGTGAG 819

Qy        974 AGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTA 1033
                || ||  ||||       ||   ||  | | |  | | |||| || | |   |  ||
Db        818 TTATTCATAATTCAGAGCCCATGTGGGATACCTTTTATCTTTTCGGCTTGTTGATTTGTA 759

Qy       1034 TGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCT 1090
              |||    ||   |||||     ||   || ||||   |  |  ||  |||| | | |
Db        758 TGTTGGTTTTGTAGAGGAAGAGTGCTCATCTGAATGTCAGAAAAACATTTTGTGGTT 702


RESULT 3
US-09-925-065A-102394
; Sequence 102394, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 102394
;   LENGTH: 534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-102394

  Query Match             3.2%;  Score 37.6;  DB 6;  Length 534;
  Best Local Similarity   53.4%;  Pred. No. 3.7;
  Matches   79;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy        159 TTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTATCC 218
              | | | ||  | |  |  || |||| ||  ||||   ||| ||||    | |  |||  |
Db        269 TCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATACCACCAGTATTAGCACTTTAAAC 328

Qy        219 CGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTTAA 278
                | || ||||| || |||||| ||||| |        ||       ||    | || ||
Db        329 ACTTGTGTTATTGTTACATGGCCTTTATTTAAATGTGTAAAATGCCTAAAGCCTTTTTCT 388

Qy        279 CCACATCGCTTCGCATGGTTACATTCTT 306
                |  |  || | |    |||||||||
Db        389 AAAACTAACTCCCAAAAAGTACATTCTT 416


RESULT 4
US-09-925-065A-849244
; Sequence 849244, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 849244
;   LENGTH: 543
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-849244

  Query Match             3.2%;  Score 37.6;  DB 6;  Length 543;
  Best Local Similarity   46.5%;  Pred. No. 3.8;
  Matches  121;  Conservative    0;  Mismatches  139;  Indels    0;  Gaps    0;

Qy        828 TTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGAGCTTTGTAGGTGGAAT 887
              || ||| | | | ||    |   ||   | |   | || | |   |||  |  | ||||
Db         28 TTATATATTTGACAAACTTCCCCAAATTGAATTCTAAATTAAGTATTTTTGACCTAGAAT 87

Qy        888 TGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGT 947
              ||   ||| | || || |||| |   |||| | |     |  ||| ||     | || |
Db         88 TGACTTTGGGATTTTCCAGTACGGCCCGTGGAGAGCCTCAAAGGATATATCCCTCATAAT 147

Qy        948 GAAGGATCCTTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTC 1007
              | ||     ||  |||  ||  ||      |       || |  ||  | ||||| |   ||
Db        148 GCAGAGATGTTAAGTTAATCAAGCTTGTTTGATGTGTTGAATTGTATGGGAAGTATTGTC 207

Qy       1008 TGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAA 1067
                || |   ||   | |   | || |  | |  |||   | || | ||  | |  |
```

```
Db        208 AAATAACAAAGTGATGTTAAATCTTTTAAAATTATATTTTATGGATATTTCAGAAATTGT 267

Qy       1068 AAACCTATCAATGTTTTCTA 1087
              |      ||| |  ||||
Db        268 ATGAAATTCATAGAAATCTA 287


RESULT 5
US-09-925-065A-902549/c
; Sequence 902549, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 902549
;   LENGTH: 605
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-902549

  Query Match             3.1%;  Score 37.4;  DB 6;  Length 605;
  Best Local Similarity   51.5%;  Pred. No. 4.5;
  Matches   86;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy       1015 TTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTA 1074
              ||| |  |  ||  | |||  | ||    | ||        |||   |   |||   | |
Db        339 TTCTTTAACTTTGTTTTTAATTTCTGCATTAAGCCAATGGTTGTTCAGAAGAATGTCGTT 280

Qy       1075 TCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTT 1134
              | | |    | ||  |  |   |  | | ||||| |||  |   || | ||| |  || |
Db        279 TAATTTCCATGTATTTGTACAATTTCCAGTGTTCTTGTTATTGATTTCTTGTTTTATTCT 220

Qy       1135 ATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAAA 1181
              |||   | |    || | ||||| |||||||||       ||||||||
Db        219 ATTGTGCCCTGAAAATATATTTGATATAGTTTTTATTTTAAAAAAAA 173


RESULT 6
US-09-925-065A-358029
; Sequence 358029, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 358029
;   LENGTH: 578
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-358029

  Query Match             3.1%;  Score 37.2;  DB 6;  Length 578;
  Best Local Similarity   50.0%;  Pred. No. 5;
  Matches   93;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy        957 TTCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTT 1016
              ||   ||| |    || | |||  | ||||   | | || |   |   |  | | | ||
Db        302 TTTTTTTTTGGTTCCTATCTTATTTCTTCTATTCATTTTGGCAACATTCTTCTTTTTTT 361

Qy       1017 CGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATC 1076
              | ||  ||| |    || |  || |    | |  ||||  |||    |||    |
Db        362 TTTTTACTTTTATAAAGTGAAAGATCAGGTCATTGAATTTGAACTTTTCTTGAACTATCC 421

Qy       1077 AATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTAT 1136
              | | ||||    ||    |||| |    || |  ||  |||||| |  ||||| ||
Db        422 ATTTTTTTAACACTTTTGCTACATCCCARAAATTTTGTATTTTGCATTTTCATTTTGAT 481

Qy       1137 TAAACT 1142
              | ||||
Db        482 TCAACT 487


RESULT 7
US-09-925-065A-106455
; Sequence 106455, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 106455
;   LENGTH: 590
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-106455

  Query Match             3.1%;  Score 37;  DB 6;  Length 590;
  Best Local Similarity   55.8%;  Pred. No. 5.7;
  Matches   92;  Conservative    0;  Mismatches   70;  Indels    3;  Gaps    1;

Qy       1005 TTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTT 1064
              | ||  || ||| || | ||||| |||| | | |||||  ||  |  ||| ||||  | |
Db        305 TCCTTTTAACTTTGTGTGGATTTTTGTTGTCTCCTATTTTCTTATAGGTGTTGATAAATC 364

Qy       1065 GAAAAACCTATCA---ATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTG 1121
              |||  |  |||  |     |||  |    ||| ||  || |||  | || | ||     ||
Db        365 TAAATATATATTACAGTAGTTCCCCTTATCCTAGGGAGATATGTTCCAAGACCCCTAGTG 424

Qy       1122 CATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTT 1166
              |||||||    | |||| | |  |         |  ||| ||| |
```

```
Db        425 CATGTGTGGGGTAATTAGAGTGAACATGTTGGGTAATTAGAGTGT 469


RESULT 8
US-09-925-065A-417521/c
; Sequence 417521, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 417521
;   LENGTH: 594
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-417521

  Query Match             3.1%;  Score 37;  DB 6;  Length 594;
  Best Local Similarity   55.8%;  Pred. No. 5.7;
  Matches   92;  Conservative    0;  Mismatches   70;  Indels    3;  Gaps    1;

Qy       1005 TTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTT 1064
              | ||  || ||| || | ||||| |||| | | ||||| ||  |  ||| ||||  | |
Db        482 TCCTTTTAACTTTGTGTKGATTTTTGTTGTCTCCTATTTTCTTATAGGTGTTGATAAATC 423

Qy       1065 GAAAAACCTATCA---ATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTG 1121
                |||  |  ||| |     |||  |    ||| ||  || |||  | ||  ||      ||
Db        422 TAAATATATATTACAGTAGTTCCCCTTATCCTAGGGAGATATGTTCCAAGACCCCTAGTG 363

Qy       1122 CATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTT 1166
              |||||||    | ||||  | |  |          |  ||| ||| |
Db        362 CATGTGTGGGGTAATTAGAGTGAACATGTTGGGTAATTAGAGTGT 318


RESULT 9
US-09-925-065A-106454
; Sequence 106454, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 106454
;   LENGTH: 599
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-106454

  Query Match             3.1%;  Score 37;  DB 6;  Length 599;
  Best Local Similarity   55.8%;  Pred. No. 5.8;
  Matches   92;  Conservative    0;  Mismatches   70;  Indels    3;  Gaps    1;

Qy       1005 TTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTT 1064
              | ||  || ||| || | ||||| |||| | | ||||| ||  |  ||| ||||  | |
Db        241 TCCTTTTAACTTTGTGTKGATTTTTGTTGTCTCCTATTTTCTTATAGGTGTTGATAAATC 300

Qy       1065 GAAAAACCTATCA---ATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTG 1121
                |||  |  ||| |     |||  |    ||| ||  || |||  | ||  ||      ||
Db        301 TAAATATATATTACAGTAGTTCCCCTTATCCTAGGGAGATATGTTCCAAGACCCCTAGTG 360

Qy       1122 CATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTT 1166
              |||||||    | ||||  | |  |          |  ||| ||| |
Db        361 CATGTGTGGGGTAATTAGAGTGAACATGTTGGGTAATTAGAGTGT 405


RESULT 10
US-09-925-065A-283413/c
; Sequence 283413, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 283413
;   LENGTH: 893
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-283413

  Query Match             3.1%;  Score 37;  DB 6;  Length 893;
  Best Local Similarity   50.9%;  Pred. No. 6.9;
  Matches   88;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        995 TGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTC 1054
              || |  ||  || |    | ||  | |||| || ||    ||| |||| |   |    |
Db        323 TGCATCAAAGTTATAACTTGTTAATATAGAATTATGGATAGTATTATTTTACTATAATTT 264

Qy       1055 TTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCC 1114
              ||| || |||  |||  | || ||   | | ||  | ||  || | | ||| |      |
Db        263 TTGTATTTTTMTAAAGGCAATAAAATGTCTTTATTTTCATTCTTGATTTTGATTCAAGTC 204

Qy       1115 TAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTT 1167
              |    |   | | | ||||  |      || | | |||  | | ||||  | ||
Db        203 TTCTGTTGTTTTTTTTTTTCCTCTTGGTCTAGCTAAAGGTATTTTATTATATT 151


RESULT 11
US-10-240-708-29
; Sequence 29, Application US/10240708
; Publication No. US20050282157A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;  TITLE OF INVENTION:  by Assessing DNA Methylation
;  FILE REFERENCE: 5013.1012
;  CURRENT APPLICATION NUMBER: US/10/240,708
;  CURRENT FILING DATE:  2002-10-03
;  PRIOR APPLICATION NUMBER: PCT/EP01/03971
;  PRIOR FILING DATE: 2001-04-06
;  PRIOR APPLICATION NUMBER: DE 10019058.8
;  PRIOR FILING DATE: 2000-04-06
;  PRIOR APPLICATION NUMBER: DE 10019173.8
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 29
;   LENGTH: 5666
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-708-29

  Query Match             3.1%;  Score 37;  DB 8;  Length 5666;
  Best Local Similarity   47.3%;  Pred. No. 16;
  Matches  112;  Conservative    0;  Mismatches  125;  Indels    0;  Gaps    0;

Qy        910 AGTTTGTGGGGTGAAAAAGCGGAGATTCGATTGATTGTGAAGGATCCTTCCGTTTCTCCG 969
Db        115 AATTATTAGGTTTATGAAGGTAAAGTTGGTTTTAGTGTTTTTTTTATTATTTTTTTAAT 174

Qy        970 GCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGT 1029
Db        175 GGTTGCGATGTTTGTTGTGGTATTTTTTATTTTATTATTGATATTGGTGTTTGTATTTTT 234

Qy       1030 GTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGC 1089
Db        235 TTTTTTTTTATTTTTATTAGTTTTGTTAGAGATTTGTTGATTTTATTGATTTTTTTAAGA 294

Qy       1090 TCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGAT 1146
Db        295 ATAAATTTTTGTTTTATTGATTTTTTTTATTGTTTTTTTATTTTTAATTTTATTGAT 351
```

RESULT 12

```
US-11-091-018-1
; Sequence 1, Application US/11091018
; Publication No. US20050287551A1
; GENERAL INFORMATION:
;  APPLICANT: Gretarsdottir, Solveig
;  APPLICANT:  Thorleifsson, Gudmar
;  APPLICANT:  Gulcher, Jeffrey R.
;  TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;
;  TITLE OF INVENTION:  METHODS OF TREATMENT
;  FILE REFERENCE: 2345.2010-016
;  CURRENT APPLICATION NUMBER: US/11/091,018
;  CURRENT FILING DATE:  2005-03-25
;  PRIOR APPLICATION NUMBER: PCT/US03/29906
;  PRIOR FILING DATE: 2003-09-25
;  PRIOR APPLICATION NUMBER: 10/255,120
;  PRIOR FILING DATE: 2002-09-25
;  PRIOR APPLICATION NUMBER: 10/419,723
;  PRIOR FILING DATE: 2003-04-18
;  PRIOR APPLICATION NUMBER: 10/650,120
;  PRIOR FILING DATE: 2003-08-27
;  PRIOR APPLICATION NUMBER: 10/067,514
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 09/811,352
;  PRIOR FILING DATE: 2001-03-19
;  NUMBER OF SEQ ID NOS: 102
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1691140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1691140)
;   OTHER INFORMATION: n=A,T,C or G
US-11-091-018-1

  Query Match             3.1%;  Score 36.4;  DB 12;  Length 1691140;
  Best Local Similarity   56.8%;  Pred. No. 2.5e+02;
  Matches   67;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy       1053 TCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGT 1112
Db     103255 TTTTGATTTTTTGGGCACCTTTAAGTTTTTTTTATATGCCATCAGGCCAACATTTATAA 103314

Qy       1113 CCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACC 1170
Db     103315 CATATGGTAGATCTGAGATGTTGTTTTACTGCCTACAAAGAATTGTTCTAGTTTTGCC 103372
```

RESULT 13

```
US-09-925-065A-881350
; Sequence 881350, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 881350
;   LENGTH: 617
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-881350

  Query Match             3.0%;  Score 36;  DB 6;  Length 617;
  Best Local Similarity   50.6%;  Pred. No. 11;
  Matches   87;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy       1006 TCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTG 1065
Db        184 TTTGGTTTTTTCTTGCTGATTTGTTTGAGTTCCTTGTAGATTCTGGATATTAGTCCTTTG 243

Qy       1066 AAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATG 1125
Db        244 TTAGACATACAGATTGTGGAGATTTTCTCCCACTCTGTGTTTTGTCTGTTTACTTTGCTG 303

Qy       1126 TGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAA 1177
Db        304 ACTATTTCTTTTGCTGTGCATGAAGCTTTTTGGTTTAGTTAACTCCCATATA 355
```

RESULT 14

US-10-750-185-46341/c

```
; Sequence 46341, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 46341
;   LENGTH: 1462
;   TYPE: DNA
;   ORGANISM: Bovine  19866880447381
US-10-750-185-46341

  Query Match             3.0%;  Score 36;  DB 8;  Length 1462;
  Best Local Similarity   48.5%;  Pred. No. 16;
  Matches   99;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        977 TTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGT 1036
              ||  ||||| || |   |   ||| | |   | |||     | || ||||  |  |
Db        937 TTTTTCCTTTACTTCCATGCTAAGTAACCGATCTTATTACTTTTTAAATTTAGGAGAAAG 878

Qy       1037 ACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGC 1096
              | | | ||     |       | || |  |  || |    || | || |    ||||||
Db        877 AATTTGATTTTTAGCAAGGGTGACATGTTTATGACTTGCTTATTTGTTTTCTACCCAAGC 818

Qy       1097 TAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTT 1156
              | ||  || || | |||||||  |            ||    ||||  ||| || || ||
Db        817 TGGCATTTTTTAAAGTCCTAAAATAATCACTCGGACTTTAAGAACTAAATCGAAGCACTT 758

Qy       1157 GTTATAGTTTTACCCCAAAAAAAA 1180
                ||  ||   | | | |||  |||
Db        757 CTTTAAGCACTTCTCAAAATTAAA 734


RESULT 15
US-10-750-623-46341/c
; Sequence 46341, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-1
;  CURRENT APPLICATION NUMBER: US/10/750,623
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 46341
;   LENGTH: 1462
;   TYPE: DNA
;   ORGANISM: Bovine  19866880447381
US-10-750-623-46341

  Query Match             3.0%;  Score 36;  DB 8;  Length 1462;
  Best Local Similarity   48.5%;  Pred. No. 16;
  Matches   99;  Conservative    0;  Mismatches  105;  Indels    0;  Gaps    0;

Qy        977 TTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGT 1036
              ||  ||||| || |   |   ||| | |   | |||     | || ||||  |  |
Db        937 TTTTTCCTTTACTTCCATGCTAAGTAACCGATCTTATTACTTTTTAAATTTAGGAGAAAG 878

Qy       1037 ACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGC 1096
              | | | ||     |       | || |  |  || |    || | || |    ||||||
Db        877 AATTTGATTTTTAGCAAGGGTGACATGTTTATGACTTGCTTATTTGTTTTCTACCCAAGC 818

Qy       1097 TAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTT 1156
              | ||  || || | |||||||  |            ||    ||||  ||| || || ||
Db        817 TGGCATTTTTTAAAGTCCTAAAATAATCACTCGGACTTTAAGAACTAAATCGAAGCACTT 758

Qy       1157 GTTATAGTTTTACCCCAAAAAAAA 1180
                ||  ||   | | | |||  |||
Db        757 CTTTAAGCACTTCTCAAAATTAAA 734


Search completed: March 20, 2006, 14:23:26
Job time : 477.098 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:59 ; Search time 236.782 Seconds
(without alignments)
8918.522 Million cell updates/sec

Title: US-10-634-548-18
Perfect score: 1188
Sequence: 1 gatacataaatcttcaacac.....ccccaaaaaaaaaaaaaaaa 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	41.6	3.5	7218	2	US-08-232-463-14	Sequence 14, Appl
	2	37.4	3.1	601	3	US-09-949-016-41156	Sequence 41156, A
	3	37.4	3.1	601	3	US-09-949-016-41157	Sequence 41157, A
c	4	37.4	3.1	161900	3	US-09-949-016-12685	Sequence 12685, A
c	5	37.4	3.1	161914	3	US-09-949-016-12906	Sequence 12906, A
c	6	36.6	3.1	100836	3	US-09-949-016-12871	Sequence 12871, A
c	7	36.6	3.1	100837	3	US-09-949-016-17063	Sequence 17063, A
	8	36.4	3.1	317366	3	US-09-949-016-16001	Sequence 16001, A
	9	35.6	3.0	404	3	US-09-640-211A-2094	Sequence 2094, Ap
	10	35.6	3.0	405	3	US-09-640-211A-354	Sequence 354, App
c	11	35.4	3.0	57638	3	US-09-949-016-17000	Sequence 17000, A
c	12	35.2	3.0	3769	3	US-09-710-279-4106	Sequence 4106, Ap
c	13	35.2	3.0	4249	3	US-09-710-279-4133	Sequence 4133, Ap
	14	35	2.9	124110	3	US-09-949-016-13353	Sequence 13353, A
	15	34.6	2.9	862	2	US-08-222-719-17	Sequence 17, Appl
	16	34.6	2.9	862	2	US-08-470-925-17	Sequence 17, Appl
	17	34.6	2.9	862	2	US-08-471-613-17	Sequence 17, Appl
	18	34.6	2.9	862	6	PCT-US93-10443-17	Sequence 17, Appl
	19	34.6	2.9	5520	3	US-10-001-887-43	Sequence 43, Appl
c	20	34.6	2.9	11168	3	US-09-434-840-5	Sequence 5, Appli
	21	34.6	2.9	50000	3	US-09-662-254B-24	Sequence 24, Appl
c	22	34.6	2.9	96690	3	US-09-949-016-17103	Sequence 17103, A
	23	34.4	2.9	832	3	US-09-621-976-2813	Sequence 2813, Ap
	24	34.4	2.9	855	3	US-09-060-726A-1	Sequence 1, Appli
c	25	34.4	2.9	856	3	US-09-060-726A-3	Sequence 3, Appli
	26	34.4	2.9	856	3	US-09-845-849A-1	Sequence 1, Appli
c	27	34.4	2.9	856	3	US-09-845-849A-3	Sequence 3, Appli
	28	34.4	2.9	4032	2	US-08-107-748-3	Sequence 3, Appli
	29	34.4	2.9	4032	2	US-08-245-809-4	Sequence 4, Appli
	30	34.4	2.9	4032	6	PCT-US92-01385-3	Sequence 3, Appli
	31	34.4	2.9	164061	3	US-09-949-016-17422	Sequence 17422, A
c	32	34.4	2.9	264665	3	US-09-949-016-13747	Sequence 13747, A
c	33	34.2	2.9	145812	3	US-09-949-016-15698	Sequence 15698, A
c	34	34.2	2.9	640681	3	US-09-790-988-1	Sequence 1, Appli
c	35	34	2.9	285	3	US-09-270-767-25614	Sequence 25614, A
c	36	34	2.9	1091	3	US-09-270-767-10251	Sequence 10251, A
c	37	34	2.9	102520	3	US-09-949-016-17367	Sequence 17367, A
c	38	34	2.9	102526	3	US-09-949-016-12448	Sequence 12448, A
	39	33.8	2.8	420	3	US-09-513-999C-32555	Sequence 32555, A
c	40	33.8	2.8	523	3	US-09-270-767-4870	Sequence 4870, Ap
c	41	33.8	2.8	523	3	US-09-270-767-20152	Sequence 20152, A
c	42	33.8	2.8	33353	3	US-09-949-016-17285	Sequence 17285, A
c	43	33.8	2.8	40000	3	US-09-780-049-18	Sequence 18, Appl
	44	33.8	2.8	89220	3	US-09-949-016-12655	Sequence 12655, A
	45	33.8	2.8	89224	3	US-09-949-016-15572	Sequence 15572, A

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s

```
US-08-232-463-14

  Query Match             3.5%;  Score 41.6;  DB 2;  Length 7218;
  Best Local Similarity   2.9%;  Pred. No. 0.055;
  Matches   11;  Conservative  212;  Mismatches  161;  Indels    0;  Gaps    0;

Qy        572 AATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACA 631
              || |  || | || ::   : : :    :: : : :::  ::::      :::  ::: :
Db       1447 AAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1388

Qy        632 TACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAACGTGATGCCACCAT 691
               :  ::  : :: :: :::::  :::   :::  :::: :::: :: : :: :  :  :
Db       1387 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1328

Qy        692 GCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCCC 751
              : : :  :: :::   :::  : :::::::    : :::::: : :::: :: :::::
Db       1327 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1268

Qy        752 ATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTTG 811
              :    : ::  : :::  : ::: : : ::: : :  ::: :: ::   :   ::    :
Db       1267 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1208

Qy        812 TTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGGA 871
               :::   : ::  ::  : : : : ::::: ::: ::::::::::::   :::: ::::
Db       1207 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1148

Qy        872 GCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGAAAAAGCGG 931
              :    : ::: ::::  : ::  : :       ::: : ::   : :::: :::::::: ::
Db       1147 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1088

Qy        932 AGATTCGATTGATTGTGAAGGATC 955
              :::   ::  ::  : :::::|||
Db       1087 RRRRRRRRRRRRRRRRRRRRRATC 1064


RESULT 2
US-09-949-016-41156
; Sequence 41156, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41156
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-41156

  Query Match             3.1%;  Score 37.4;  DB 3;  Length 601;
  Best Local Similarity   55.0%;  Pred. No. 0.29;
  Matches   94;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1011 TATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATAT-TTGAAAA 1069
              ||| || |||||  ||| |||  | |   ||  | | ||  ||      | | ||| |||
Db        331 TATTTTTGTCTAACTTTATGTATTTTTTAATCGTAAAAGACGTTAAATTTTTGTTGTAAA 390

Qy       1070 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 1129
              | |||||||| |||| |   ||| | | |   |||||   | || || |||   | |  |
Db        391 AACTATCAATATTTTATTTTTCCTATCAAAAGATTGTATTTTTCTTATGTTTAGTTTTAA 450

Qy       1130 TTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
                |      ||| | ||        ||  |||    |||||||||   | ||
Db        451 CTAGATGAAAATTTGAGGTTCTGTTTATTTAAGCCTTTACCCCATCTATAA 501


RESULT 3
US-09-949-016-41157
; Sequence 41157, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41157
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-41157

  Query Match             3.1%;  Score 37.4;  DB 3;  Length 601;
  Best Local Similarity   55.0%;  Pred. No. 0.29;
  Matches   94;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1011 TATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATAT-TTGAAAA 1069
              ||| || |||||  ||| |||  | |   ||  | | ||  ||      | | ||| |||
Db        349 TATTTTTGTCTAACTTTATGTATTTTTTAATCGTAAAAGACGTTAAATTTTTGTTGTAAA 408

Qy       1070 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 1129
              | |||||||| |||| |   ||| | | |   |||||   | || || |||   | |  |
Db        409 AACTATCAATATTTTATTTTCCTATCAAAAGATTGTATTTTTCTTATGTTTAGTTTTAA 468

Qy       1130 TTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
                |      ||| | ||        ||  |||    |||||||||   | ||
Db        469 CTAGATGAAAATTTGAGGTTCTGTTTATTTAAGCCTTTACCCCATCTATAA 519


RESULT 4
US-09-949-016-12685/c
; Sequence 12685, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12685
;   LENGTH: 161900
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(161900)
;   OTHER INFORMATION: n = A,T,C or G
```

```
US-09-949-016-12685

  Query Match             3.1%;  Score 37.4;  DB 3;  Length 161900;
  Best Local Similarity   55.0%;  Pred. No. 7.2;
  Matches   94;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1011 TATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATAT-TTGAAAA 1069
              ||| || |||||  ||| |||  | |   ||  | | ||  ||       | | ||| |||
Db      81619 TATTTTTGTCTAACTTTATGTATTTTTTAATCGTAAAAGACGTTAAATTTTTGTTGTAAA 81560

Qy       1070 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 1129
              | ||||||||| |||| |   ||| | | |   |||||   | || || |||   | | |
Db      81559 AACTATCAATATTTTATTTTTCCTATCAAAAGATTGTATTTTTCTTATGTTTAGTTTTAA 81500

Qy       1130 TTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
              |      ||| | ||          ||  |||    |||||||||   | ||
Db      81499 CTAGATGAAAATTTGAGGTTCTGTTTATTTAAGCCTTTACCCCATCTATAA 81449


RESULT 5
US-09-949-016-12906/c
; Sequence 12906, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12906
;   LENGTH: 161914
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(161914)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12906

  Query Match             3.1%;  Score 37.4;  DB 3;  Length 161914;
  Best Local Similarity   55.0%;  Pred. No. 7.2;
  Matches   94;  Conservative    0;  Mismatches   76;  Indels    1;  Gaps    1;

Qy       1011 TATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATAT-TTGAAAA 1069
              ||| || |||||  ||| |||  | |   ||  | | ||  ||       | | ||| |||
Db      81629 TATTTTTGTCTAACTTTATGTATTTTTTAATCGTAAAAGACGTTAAATTTTTGTTGTAAA 81570

Qy       1070 ACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTA 1129
              | ||||||||| |||| |   ||| | | |   |||||   | || || |||   | | |
Db      81569 AACTATCAATATTTTATTTTTCCTATCAAAAGATTGTATTTTTCTTATGTTTAGTTTTAA 81510

Qy       1130 TTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
              |      ||| | ||          ||  |||    |||||||||   | ||
Db      81509 CTAGATGAAAATTTGAGGTTCTGTTTATTTAAGCCTTTACCCCATCTATAA 81459


RESULT 6
US-09-949-016-12871/c
; Sequence 12871, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12871
;   LENGTH: 100836
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(100836)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12871

  Query Match             3.1%;  Score 36.6;  DB 3;  Length 100836;
  Best Local Similarity   48.3%;  Pred. No. 9.9;
  Matches  102;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy        958 TCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTC 1017
              ||   ||| ||      | ||| ||||      |||  ||  ||      ||     |   |
Db      35075 TCAACTTCCCCAATTTACCTTAATCCAATCTCTGCCTTCCAATTCCTTTGGACTCAGATT 35016

Qy       1018 GTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCA 1077
              |||||| |  ||| | || |   |    |    |||| |||| ||  | ||  || |
Db      35015 CAGAAGATTTTTTATATTTTCTGTATTTTAATTATTCTTTAATAATTTCACAATATAACC 34956

Qy       1078 ATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATT 1137
              | || |||  | || |     |  ||  | |||   |   |   |  |||    |||
Db      34955 TTTTTATCTGACCCCTATGCTCCCTTTAATTATGAAATTTTTGAGAAATGTGACATTAGA 34896

Qy       1138 AAACTCGATCAAAACATTTGTTATAGTTTTA 1168
              | |||   ||   |    |||||||| | |||
Db      34895 ACACTTCTTCTCTAACAATGTTATATTATTA 34865


RESULT 7
US-09-949-016-17063/c
; Sequence 17063, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17063
;   LENGTH: 100837
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(100837)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17063

  Query Match             3.1%;  Score 36.6;  DB 3;  Length 100837;
  Best Local Similarity   48.3%;  Pred. No. 9.9;
```

```
  Matches  102;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy        958 TCCGTTTCTCCGGCCAAGCTTGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTC 1017
              ||   ||| ||     | ||| ||||     |||  ||  ||     ||    |   |
Db      35075 TCAACTTCCCCAATTTACCTTAATCCAATCTCTGCCTTCCAATTCCTTTGGACTCAGATT 35016

Qy       1018 GTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCA 1077
              |||||| |  ||| | || |   |    |    |||| |||| ||  | ||  || |
Db      35015 CAGAAGATTTTTTATATTTTCTGTATTTTAATTATTCTTTAATAATTTCACAATATAACC 34956

Qy       1078 ATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATT 1137
              | || |||  | || |     |  ||  | |||   |   |   |  |||    |||
Db      34955 TTTTTATCTGACCCCTATGCTCCCTTTAATTATGAAATTTTTGAGAAATGTGACATTAGA 34896

Qy       1138 AAACTCGATCAAAACATTTGTTATAGTTTTA 1168
              | |||   ||   |   ||||||| | |||
Db      34895 ACACTTCTTCTCTAACAATGTTATATTATTA 34865


RESULT 8
US-09-949-016-16001
; Sequence 16001, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16001
;   LENGTH: 317366
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(317366)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16001

  Query Match             3.1%;  Score 36.4;  DB 3;  Length 317366;
  Best Local Similarity   49.5%;  Pred. No. 22;
  Matches   94;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy        978 TGATCCTTCACCTGAGTTGGAAGAAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTA 1037
              | || |   |  |   | ||||  |  |||       |   | || | ||||||| | ||
Db     151655 TTATTCAATAAATTTTTAGGAATTATATTCAATGTATTCATTTTATAATTGTGTTCTTTA 151714

Qy       1038 CTATTATCAGAGGGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCT 1097
              || |||       | ||   | ||  ||||   || |||| || | |   |    |
Db     151715 ATAATATATTTTTAATGTTATTTTTTACAAAACTTTACCAATTTTCTTTTTTTTTTTGAA 151774

Qy       1098 AGCTATTGTTCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTG 1157
              |  | |  || | || |   | | |||||| |  ||| ||| |   |   |   ||| |
Db     151775 ACTTTTATTTTAGGTTCAGGGGTACATGTGCAGGTTTGTTATATAGGTATACTTCATGTC 151834

Qy       1158 TTATAGTTTT 1167
              |   |||||
Db     151835 ATGGGGTTTT 151844


RESULT 9
US-09-640-211A-2094
; Sequence 2094, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2094
;   LENGTH: 404
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-2094

  Query Match             3.0%;  Score 35.6;  DB 3;  Length 404;
  Best Local Similarity   51.2%;  Pred. No. 0.86;
  Matches   83;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        559 GCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCT 618
              | |||||  | |||      |    ||  || |   || ||       |  ||||  |||  |
Db        219 GAAGGAAGTGACACGAGTACATCAAAAGAGCAAGATGTCCAGTGTTTAAAACGAGAAAGT 278

Qy        619 TTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAAC 678
              |  ||||  |  | ||| | |  | | | ||   |   ||  || || | | |
Db        279 GCGAATATGGAAGAAAGGATTGAAATTCTTGAATCCATGCAAAGAAAGATGTTGGGCGAG 338

Qy        679 GTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              | | || || |||| |||   | ||| || || ||   ||||
Db        339 GAGCTGGCATCATGTGCATTGAAGGATTTGAATCAGTTGGAG 380


RESULT 10
US-09-640-211A-354
; Sequence 354, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 354
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-354

  Query Match             3.0%;  Score 35.6;  DB 3;  Length 405;
  Best Local Similarity   51.2%;  Pred. No. 0.86;
  Matches   83;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        559 GCAGGAACTAACAAATACATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCT 618
              | |||||  | |||      |    ||  || |   || ||       |  ||||  |||  |
Db        219 GAAGGAAGTGACACGAGTACATCAAAAGAGCAAGATGTCCAGTGTTTAAAACGAGAAAGT 278

Qy        619 TTCGAGCTGGACATACCGGTTGCAGTGGTGGGAACCGGACTCGGACCGAAGTGGAACAAC 678
              |  ||||   |  | ||| | |  | | | ||   |   ||  || || | | |
Db        279 GCGAATATGGAAGAAAGGATTGAAATTCTTGAATCCATGCAAAGAAAGATGTTGGGCGAG 338

Qy        679 GTGATGCCACCATGCGCACCAACGGACTTAAACCATGAGGAG 720
              | | || || |||| |||   | ||| || || ||   ||||
Db        339 GAGCTGGCATCATGTGCATTGAAGGATTTGAATCAGTTGGAG 380
```

```
RESULT 11
US-09-949-016-17000/c
; Sequence 17000, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17000
;   LENGTH: 57638
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(57638)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17000

  Query Match             3.0%;  Score 35.4;  DB 3;  Length 57638;
  Best Local Similarity   57.8%;  Pred. No. 17;
  Matches   63;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy       1050 GGGTCTTGAATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCA 1109
              |||  | | | |       ||||  |  | |  ||| ||  || ||  || ||  ||   |
Db      19796 GGGGATGGGACACAGTTTAAACACAAAATTCATTTATATTCAAACATACCTTATGCATA 19737

Qy       1110 TGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGT 1158
                 ||| | ||| ||||||| |||||| ||||  |  || |  |||||
Db      19736 CAGCCTGAATTGTATGTGTAATTTTATACAACTTAAAAAATAATTTTGT 19688

RESULT 12
US-09-710-279-4106/c
; Sequence 4106, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4106
;   LENGTH: 3769
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4106

  Query Match             3.0%;  Score 35.2;  DB 3;  Length 3769;
  Best Local Similarity   49.5%;  Pred. No. 4.2;
  Matches   91;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy       1001 AAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAAT 1060
               || ||    ||||    |  | |||  |||           ||||| || |    | || ||
Db       3181 AATGTTTGCATATCCAATTTCAAATTGATGTATGTGTACATTATAAGCGAATTGTTAAAA 3122

Qy       1061 ATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTT 1120
               | |  ||       |     |  ||| ||   |   |   |     |    | ||    | |||
Db       3121 AGTGAAATTTGTTTCTTTTTATTTTTAAAACATTGTAATTGAAAACTAGTTTATAATGTT 3062

Qy       1121 GCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
               | || | | ||  |||| || | ||| || | | |||| || ||| |  |   | | | ||
Db       3061 GAATCTTTTTAGTATTTAAATAGATTAATAAAATTGTATTATTTTAAGCTTAGACATAA 3002

Qy       1181 AAAA 1184
                |||
Db       3001 CAAA 2998

RESULT 13
US-09-710-279-4133/c
; Sequence 4133, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/09/710,279
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4133
;   LENGTH: 4249
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4133

  Query Match             3.0%;  Score 35.2;  DB 3;  Length 4249;
  Best Local Similarity   49.5%;  Pred. No. 4.5;
  Matches   91;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy       1001 AAGCTTCTGGTATCTTCGTCTAGATTTGTGTTATGTACTATTATCAGAGGGGTCTTGAAT 1060
               || ||    ||||    |  | |||  |||           ||||| || |    | || ||
Db       1763 AATGTTTGCATATCCAATTTCAAATTGATGTATGTGTACATTATAAGCGAATTGTTAAAA 1704

Qy       1061 ATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAAGTT 1120
               | |  ||       |     |  ||| ||   |   |   |     |    | ||    | |||
Db       1703 AGTGAAATTTGTTTCTTTTTATTTTTAAAACATTGTAATTGAAAACTAGTTTATAATGTT 1644

Qy       1121 GCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAAAAA 1180
               | || | | ||  |||| || | ||| || | | |||| || ||| |  |   | | | ||
Db       1643 GAATCTTTTTAGTATTTAAATAGATTAATAAAATTGTATTATTTTAAGCTTAGACATAA 1584

Qy       1181 AAAA 1184
                |||
Db       1583 CAAA 1580

RESULT 14
US-09-949-016-13353
; Sequence 13353, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
```

```
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13353
;   LENGTH: 124110
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13353

  Query Match             2.9%;  Score 35;  DB 3;  Length 124110;
  Best Local Similarity   54.2%;  Pred. No. 36;
  Matches   71;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy       1058 AATATTTGAAAAACCTATCAATGTTTTCTAGCTCCAAGCTAGCTATTGTTCATGTCCTAA 1117
              |||  ||     |  |||||||| | | || |  ||  |||||  |||| |  |   |
Db     100910 AATGATTTCTGCAGTCATCAATGTATACAAGTTAAAATTTAGCTCCTGTTAAATTAACAC 100969

Qy       1118 GTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTTTACCCCAAAAA 1177
              |        |||    || |||        ||| | || ||   ||| ||| |||| |
Db     100970 TTCATTAAAATATCAAAATAAAAGCTTCGTCAAATACTTTTTTATGTTGTACTCCAAGAC 101029

Qy       1178 AAAAAAAAAAA 1188
              | |||||||||
Db     101030 AGAAAAAAAAA 101040


RESULT 15
US-08-222-719-17
; Sequence 17, Application US/08222719
; Patent No. 5846711
;  GENERAL INFORMATION:
;    APPLICANT:  David D. Moore
;    APPLICANT:  Jae Woon Lee
;    TITLE OF INVENTION:  NUCLEAR HORMONE RECEPTOR-
;    TITLE OF INVENTION:  INTERACTING POLYPEPTIDES AND
;    TITLE OF INVENTION:  RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/222,719
;      FILING DATE:  04-April-1994
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,136
;      FILING DATE:  30-October-1992
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/229001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  862
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-222-719-17

  Query Match             2.9%;  Score 34.6;  DB 2;  Length 862;
  Best Local Similarity   64.2%;  Pred. No. 2.8;
  Matches   52;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       1107 TCATGTCCTAAGTTGCATGTGTATTTTTATTAAACTCGATCAAAACATTTGTTATAGTTT 1166
              ||||||  |   | |||  | |  |||   | || |  || |||| |||  | ||  |||
Db        782 TCATGTATTTGTTCGCAATTATGCTTTCTCTGAATTTAATAAAAATATTCCTAATGCTTT 841

Qy       1167 TACCCCAAAAAAAAAAAAAAA 1187
              ||   ||||||||||||||||
Db        842 TAGAAAAAAAAAAAAAAAAAA 862


Search completed: March 20, 2006, 14:05:53
Job time : 240.782 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:54:38 ; Search time 4014.27 Seconds
(without alignments)
16071.974 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1135	100.0	1135	6	E63072	E63072 DNA encodin
	2	1135	100.0	1135	15	AF134302	AF134302 Arabidops
	3	998.2	87.9	1025	15	BT002898	BT002898 Arabidops
	4	803	70.7	87286	15	AB026651	AB026651 Arabidops
	5	745	65.6	1104	15	AF337545	AF337545 Brassica
	6	390	34.4	884	15	AF337546	AF337546 Brassica
	7	367.4	32.4	1216	6	AX411603	AX411603 Sequence
	8	338.6	29.8	1159	15	BT013523	BT013523 Lycopersi
	9	281.6	24.8	128855	14	AP007972	AP007972 Lotus cor
	10	256.8	22.6	1444	6	AX411613	AX411613 Sequence
	11	247.4	21.8	1157	15	AY292526	AY292526 Ginkgo bi
	12	237.4	20.9	987	6	AX411601	AX411601 Sequence
	13	221.6	19.5	242	15	ATH524543	AJ524543 Arabidops
	14	201.6	17.8	1302	6	AX411605	AX411605 Sequence
	15	183.4	16.2	1115	15	AF337544	AF337544 Brassica
	16	163.4	14.4	975	6	AX412268	AX412268 Sequence
	17	163.4	14.4	975	6	AX412469	AX412469 Sequence
	18	163.4	14.4	975	6	AX505463	AX505463 Sequence
	19	163.4	14.4	1139	15	BT000309	BT000309 Arabidops
	20	163.4	14.4	1188	6	E63071	E63071 DNA encodin
	21	163.4	14.4	1188	15	AF021244	AF021244 Arabidops
	22	163.4	14.4	1194	15	AY099815	AY099815 Arabidops
	23	151	13.3	1174	6	AX411607	AX411607 Sequence
	24	148.2	13.1	1242	6	AX411617	AX411617 Sequence
	25	148.2	13.1	1242	15	BT009214	BT009214 Triticum
	26	145.4	12.8	1104	6	AX411609	AX411609 Sequence
	27	142.8	12.6	1125	6	AX411611	AX411611 Sequence
	28	136	12.0	1157	15	AY089091	AY089091 Arabidops
	29	121.2	10.7	992	15	AB056126	AB056126 Citrus un
	30	118	10.4	1092	15	AF160869	AF160869 Citrus si
	31	114.4	10.1	90341	15	AC024609	AC024609 Arabidops
	32	114.4	10.1	119942	15	AC007797	AC007797 Arabidops
c	33	109	9.6	110000	15	AP008216_142	Continuation (143
c	34	109	9.6	142737	15	AC027658	AC027658 Oryza sat
c	35	109	9.6	303145	15	AE017095	AE017095 Oryza sat
	36	93.6	8.2	1340	6	E63070	E63070 DNA encodin
	37	93.6	8.2	1340	15	AB025025	AB025025 Chenopodi
	38	93.6	8.2	1340	15	AF134301	AF134301 Chenopodi
	39	88.4	7.8	1156	6	AX411625	AX411625 Sequence
	40	88.4	7.8	1156	15	BT008923	BT008923 Triticum
	41	71.4	6.3	916	6	AX411615	AX411615 Sequence
	42	71.4	6.3	916	15	BT009115	BT009115 Triticum
c	43	66.6	5.9	110000	15	AP008213_280	Continuation (281
c	44	66.6	5.9	162545	15	AP004274	AP004274 Oryza sat
	45	64.8	5.7	1300	6	E63080	E63080 DNA encodin

ALIGNMENTS

RESULT 1
E63072
LOCUS E63072 1135 bp DNA linear PAT 27-AUG-2002
DEFINITION DNA encoding chlorophylase and plant transformed by it.
ACCESSION E63072
VERSION E63072.1 GI:22553555
KEYWORDS JP 2001086990-A/3.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1135)
AUTHORS Tsuchiya,T., Ota,H., Takamiya,K., Harada,S. and Nakat,K.
TITLE DNA encoding chlorophylase and plant transformed by it
JOURNAL Patent: JP 2001086990-A 3 03-APR-2001;
KAGOME CO LTD
COMMENT OS Arabidopsis thaliana (thale cress)
PN JP 2001086990-A/3
PD 03-APR-2001
PF 20-SEP-1999 JP 1999266181
PI TORU TSUCHIYA,HIROYUKI OTA,KENICHIRO TAKAMIYA,SATOSHI HARADA,
PI KENGO NAKATA
PC C12N15/09,A01H5/00,C12N5/10,C12N9/16//(C12N9/16,C12R1:19), PC C12N15/00,
PC C12N5/00
CC
FH Key Location/Qualifiers
FT CDS (43). .(996).
FEATURES Location/Qualifiers
source 1. .1135
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
ORIGIN

Query Match 100.0%; Score 1135; DB 6; Length 1135;
Best Local Similarity 100.0%; Pred. No. 5.7e-271;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60

```
Db          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180

Qy        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240

Qy        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300

Qy        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360

Qy        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420

Qy        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480

Qy        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540

Qy        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600

Qy        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660

Qy        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720

Qy        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780

Qy        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840

Qy        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900

Qy        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960

Qy        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAA 1135

RESULT 2
AF134302
LOCUS       AF134302                1135 bp    mRNA    linear   PLN 22-JAN-2000
DEFINITION  Arabidopsis thaliana AtCLH2 (AtCLH2) mRNA, complete cds.
ACCESSION   AF134302
VERSION     AF134302.1  GI:6729676
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1135)
  AUTHORS   Tsuchiya,T., Ohta,H., Okawa,K., Iwamatsu,A., Shimada,H., Masuda,T.
            and Takamiya,K.
  TITLE     Cloning of chlorophyllase, the key enzyme in chlorophyll
            degradation: finding of a lipase motif and the induction by methyl
            jasmonate
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (26), 15362-15367 (1999)
   PUBMED   10611389
REFERENCE   2  (bases 1 to 1135)
  AUTHORS   Tsuchiya,T., Ohta,H., Okawa,K., Iwamatsu,A., Shimada,H., Masuda,T.
            and Takamiya,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-1999) Department of Biological Sciences, Tokyo
            Institute of Technology, Yokohama 226-8501, Japan
FEATURES             Location/Qualifiers
     source          1. .1135
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
     gene            1. .1135
                     /gene="AtCLH2"
     CDS             43. .999
                     /gene="AtCLH2"
                     /EC_number="3.1.1.14"
                     /note="chlorophyllase 2; hydrolase; includes a p-loop
                     motif and lipase/serine active site motif"
                     /codon_start=1
                     /product="AtCLH2"
                     /protein_id="AAF27046.1"
                     /db_xref="GI:6729677"
                     /translation="MSSSSSRNAFEDGKYKSNLLTLDSSSRCCKITPSSRASPSPPKQ
                     LLVATPVEEGDYPVVMLLHGYLLYNSFYSQLMLHVSSHGFILIAPQLYSIAGPDTMDE
                     IKSTAEIMDWLSVGLNHFLPAQVTPNLSKFALSGHSRGGKTAFAVALKKFGYSSNLKI
                     STLIGIDPVDGTGKGKQTPPPVLAYLPNSFDLDKTPILVIGSGLGETARNPLFPPCAP
                     PGVNHREFFRECQGPAWHFVAKDYGHLDMLDDDTKGIRGKSSYCLCKNGEERRPMRRF
                     VGGLVVSFLKAYLEGDDRELVKIKDGCHEDVPVEIQEFEVIM"
ORIGIN

  Query Match           100.0%;  Score 1135;  DB 15;  Length 1135;
  Best Local Similarity  100.0%;  Pred. No. 5.7e-271;
  Matches 1135;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180

Qy        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
```

```
Qy        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300

Qy        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360

Qy        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420

Qy        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480

Qy        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540

Qy        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600

Qy        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660

Qy        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720

Qy        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780

Qy        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840

Qy        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900

Qy        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960

Qy        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
```

```
RESULT 3
BT002898
LOCUS       BT002898               1025 bp    mRNA    linear   PLN 15-JAN-2003
DEFINITION  Arabidopsis thaliana clone RAFL15-48-F08 (R20874) putative AtCLH2
            protein (At5g43860) mRNA, complete cds.
ACCESSION   BT002898
VERSION     BT002898.1  GI:27754533
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1025)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,
            Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
            Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
            Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1025)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
            Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
            Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
            Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M.,
            Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G.,
            Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,
            Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,
            Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.

            Annotation based on July 2002 version of the Arabidopsis genome
            submitted to Genbank.
FEATURES                Location/Qualifiers
     source             1. .1025
                        /organism="Arabidopsis thaliana"
                        /mol_type="mRNA"
                        /db_xref="taxon:3702"
                        /chromosome="5"
                        /clone="RAFL15-48-F08 (R20874)"
                        /ecotype="Columbia"
                        /note="This clone is in a modified pBluescript vector2
                        (lambda PS) as a BamHI/XhoI insert."
     gene               <1. .1025
                        /gene="At5g43860"
     CDS                <1. .883
                        /gene="At5g43860"
                        /codon_start=2
                        /evidence=experimental
                        /product="putative AtCLH2 protein"
                        /protein_id="AAO22714.1"
                        /db_xref="GI:27754534"
                        /translation="SCCCKITPSSRASPSPPKQLLVATPVEEGDYAVVMLLHGYLLYN
                        SFYSQLMLHVSSHGFILIAPQLYSIAGPDTMDEIKSTAEIMDWLSVGLNHFLPAQVTP
                        NLSKFALSGHSRGGKTAFAVALKKFGYSSNLKISTLIGIDPVDGTGKGKQTPPPVLAY
                        LPNSFDLDKTPILVIGSGLGETARNPLFPPCAPPGVNHREFFRECQGPAWHFVAKDYG
                        HLDMLDDDTKGIRGKSSYCLCKNGEERRPMRRFVGGLVVSFLKAYLEGDDRELVKIKD
                        GCHEDVPVEIQEFEVIM"
     misc_difference    5
                        /gene="At5g43860"
                        /note="compared to genomic sequence resulting in an amino
                        acid sequence difference"
                        /replace="c"
     misc_difference    95
```

```
                     /gene="At5g43860"
                     /note="compared to genomic sequence resulting in an amino
                     acid sequence difference"
                     /replace="c"
     3'UTR           884. .1025
                     /gene="At5g43860"
     misc_difference 1010
                     /gene="At5g43860"
                     /note="compared to genomic sequence"
                     /replace="t"
ORIGIN

  Query Match             87.9%;  Score 998.2;  DB 15;  Length 1025;
  Best Local Similarity   98.7%;  Pred. No. 5.8e-237;
  Matches 1006;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy        117 ATCTCGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCT 176
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCT 60

Qy        177 GTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTA 236
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         61 GTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATGCGGTGGTGATGCTCCTCCATGGTTA 120

Qy        237 CCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCAT 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCAT 180

Qy        297 CCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATC 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATC 240

Qy        357 AACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGT 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGT 300

Qy        417 AACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTT 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTT 360

Qy        477 TGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGG 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGG 420

Qy        537 TATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTA 596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTA 480

Qy        597 CCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGG 656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGG 540

Qy        657 TGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGA 716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGA 600

Qy        717 GTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTT 776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTT 660

Qy        777 GGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAA 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAA 720

Qy        837 GAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTT 896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTT 780

Qy        897 GAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGA 956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGA 840

Qy        957 GGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTA 1016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTA 900

Qy       1017 GGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTT 1076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTT 960

Qy       1077 ATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||||||||||||||||||| | |  ||    |  ||||||||
Db        961 ATATTGTACAACTCTTTAAGTCACCTCTTTGCTTATGATATTAACCCGACCAAAAAAAA 1019


RESULT 4
AB026651
LOCUS       AB026651               87286 bp    DNA     linear   PLN 14-FEB-2004
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MQD19.
ACCESSION   AB026651 BA000015
VERSION     AB026651.1  GI:4757407
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Kaneko,T., Katoh,T., Asamizu,E., Sato,S., Nakamura,Y., Kotani,H.
            and Tabata,S.
  TITLE     Structural analysis of Arabidopsis thaliana chromosome 5. XI
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 87286)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MQD19
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers.  Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
            http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
            Genes encoding tRNAs are predicted by tRNAscan-SE
            (Sean Eddy, Washington University School of Medicine, St. Louis,
            http://genome.wustl.edu/eddy/tRNAscan-SE/).
            This sequence may not be the entire insert of this clone. It may be
            shorter because we remove overlaps between neighboring submissions.
            The 5' clone is MQO24 and the 3' clone is F6B6.
FEATURES             Location/Qualifiers
     source          1. .87286
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="MQD19"
                     /clone_lib="Mitsui P1"
                     /ecotype="Columbia"
     CDS             1164. .2159
                     /note="gene_id:MQD19.2"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="steroid sulfotransferase-like"
                     /protein_id="BAB11296.1"
                     /db_xref="GI:10177937"
                     /translation="MAGKSDLPVIARDDNLSEETKTLISSLPTYQDSHVKLCKYQGCW
```

YYHNTLQAVINYQRNFQPQDTDIILASFPKSGTTWLKALSVAIVERSKQPFDDDPLTH
PLLSDNPHGIVPFFEFDMYLKTSTPDLTKFSTSSPRLFSTHMPLHTFKEGLKGSPCKV
VYMCRNIKDVLISDWHFRSKYSNNEVSRSTLESMFESFCGGVSFYGPFWDHALSYWRG
SLENPKHVLFMRYEEMKTEPCVQVKRLAEFLGFPFTKEEEDSGSISKLLELCSLGNLS
GLEVNKTGKTWMNYDYKSYFRKGEVGDWKNHLTPEMENKIDMIIEEKLKGSDLKF"

CDS join(5356. .5534,5627. .5929,6019. .6097)
/note="gene_id:MQD19.3
sp|P33077"
/codon_start=1
/evidence=not_experimental
/product="auxin-induced protein AUX2-11"
/protein_id="BAB11297.1"
/db_xref="GI:10177938"
/translation="MEKVDVYDELVNLKATELRLGLPGTEETVSCGKSNKRVLPEATE
KEIESTGKTETASPPKAQIVGWPPVRSYRKNNVQTKKSESEGQGNYVKVSMDGAPYLR
KIDLTMYKQYPELMKSLENMFKFSVGEYFEREGYKGSDFVPTYEDKDGDWMLVGDVPW
EMFVSSCKRLRIMKGSEVKGLGCGGL"

CDS complement(join(7143. .7430,7666. .7870,7958. .8028,
8118. .8253,8345. .8424,8530. .8643,8737. .8818,8906. .8994,
9168. .9218,9339. .9448,9635. .9707,9969. .10024,10108. .10198,
10456. .10542,10681. .10773,10857. .10976,11286. .11414))
/note="unnamed protein product; dbj|BAA91806.1
gene_id:MQD19.4
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11298.1"
/db_xref="GI:10177939"
/translation="MDSNFKWLLFAILISLTFSGFVLHHGVLAESVKPDEAKQLRDEV
RGMFYHAFDGYMNNAFPLDELRPLSCQGEDTLGGYALTLIDSLDTLALLGDRERFTSS
VEWIGKNLQFNINKTVSVFETTIRVLGGLLSAHLIASDYATGMRIPSYNNELLVLAEN
LARRMLPAFDTPTGIPFGSVNLMYGVDKHESKITSTAGGGTLSLEFGVLSRLTNDPVF
EQVAKNAVRGLWARRSNLDLVGAHINVFTGEWTQKDAGIGTSIDSFYEYLLKAYILFG
DEEYLYIFQEAYRSAMQYLHKDPWYVEVNMDSAAIVWPVFNSLQAFWPGLQVLAGDVD
PAIRTHTAPFSVWKRYGFTPEGFNLATLSVQYGQKSYPLRPELIESTYWLYKATRDPR
YLDAGRDFVASLQYGAKCPCGYCHITDVELHKQEDHMESFFLAETVKYLWLLFDLAVD
SDNLVDNGPYKYIFSTEGHLLPITPQISLAREHCSYFGGYCPSNSTKLEQEVLGEDSS
NDDHSNDYPYHESFPVTGLIKGLCPGLTHAQKYGPSYVLPEKTDREDVNQPKPVVTSS
SIVLISDQTVEKRPQEEEGFTSQSEPIMTISGGSSNDQTGQELTLLESETDDQRSYSS
"

CDS complement(join(12391. .12699,12792. .12855,12945. .13059,
13177. .13315,13415. .13542,13630. .13759,13977. .14129))
/note="unnamed protein product; gene_id:MQD19.5
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11299.1"
/db_xref="GI:10177940"
/translation="MAHGGYAKRRDSEPKQTAGSSRRSKVLRVEKKPKTVSIKNQMRS
VERFLRKDLPPEVRESLVEKLEDLKKQQDAHARLAVERKIFFRNRKIKPFERRKIERS
IRRLEKLQRTSSAHVGDVDIAEQLSKLKEDLEYVRFFPKNEKYVSLFTGAEDSEVIEK
RGKMRQQIKANIIVAAASGKELEETGSEDDGLLDLSDDDFFDKGSSSDEADADDELTD
KSAKEAASSASGRATSGMSSDERNQKQNSDRALMPPPQARFESNSRKNLYVQRNEMPS
SSRNNTSNRRSESSYNARASAANSYSSQSSNLSSNSDAHKPKRKRRPKKKKQQVGLLT
YELICRKQIFE"

CDS 15158. .17704
/note="gene_id:MQD19.6"
/codon_start=1
/evidence=not_experimental
/product="disease resistance protein"
/protein_id="BAB11300.1"
/db_xref="GI:10177941"
/translation="MVDWLSLLPWNKIFTAACGCFLSDSNYIHLMESNLDALQKTMEE
LKNGRDDLLARVSIEEDKGLQRLALVNGWLSRVQIVESEFKDLLEAMSIETGRLCLFG
YCSEDCISSYNYGGKVMKNLEEVKELLSKKNFEVVAQKIIPKAEKKHIQTTVGLDTMV
GIAWESLIDDEIRTLGLYGMGGIGKTTLLESLNNKFVELESEFDVVIWVVVSKDFQLE
GIQDQILGRLRPDKEWERETESKKASLINNNLKRKKFVLLLDDLWSEVDLIKIGVPPP
SRENGSKIVFTTRSKEVCKHMKADKQIKVDCLSPDEAWELFRLTVGDIILRSHQDIPA
LARIVAAKCHGLPLALNVIGKAMVCKETVQEWRHAINVLNSPGHKFPGMEERILPILK
FSYDSLKNGEIKLCFLYCSLFPEDFEIEKDKLIEYWICEGYINPNRYEDGGTNQGYDI
IGLLVRAHLLIECELTDKVKMHDVIREMALWINSDFGNQQETICVKSGAHVRLIPNDI
SWEIVRQMSLISTQVEKIACSPNCPNLSTLLLPYNKLVDISVGFFLFMPKLVVLDLST
NWSLIELPEEISNLGSLQYLNLSLTGIKSLPVGLKKLRKLIYLNLEFTNVLESLVGIA
TTLPNLQVLKLFYSLFCVDDIIMEELQRLKHLKILTATIEDAMILERVQGVDRLASSI
RGLCLRNMSAPRVILNSVALGGLQQLGIVSCNISEIEIDWLSKERRDHRSTSSPGFKQ
LASITVIGLVGPRDLSWLLFAQNLKDIQVQYSPTIEEIINKQKGMSITKVHRDIVVPF
GKLESLHLYQLAELTEICWNYQTLPNLRESYVNYCPKLLEDIANFPKLKG"

CDS 20901. .23489
/note="gene_id:MQD19.7"
/codon_start=1
/evidence=not_experimental
/product="disease resistance protein"
/protein_id="BAB11301.1"
/db_xref="GI:10177942"
/translation="MLGWLVIPWNQIFTAACGCFLSDRNYIHMMESNLDALQKTMEEL
KNGRDDLLGRVSIEEDKGLQRLAQVNGWLSRVQIVESEFKDLLEAMSIETGRLCLLGY
CSEDCISSYNYGEKVSKMLEEVKELLSKKDFRMVAQEIIHKVEKKLIQTTVGLDKLVE
MAWSSLMNDEIGTLGLYGMGGVGKTTLLESLNNKFVELESEFDVVIWVVVSKDFQFEG
IQDQILGRLRSDKEWERETESKKASLIYNNLERKKFVLLLDDLWSEVDMTKIGVPPPT
RENGSKIVFTTRSTEVCKHMKADKQIKVACLSPDEAWELFRLTVGDIILRSHQDIPAL
ARIVAAKCHGLPLALNVIGKAMSCKETIQEWSHAINVLNSAGHEFPGMEERILPILKF
SYDSLKNGEIKLCFLYCSLFPEDSEIPKEKWIEYWICEGFINPNRYEDGGTNHGYDII
GLLVRAHLLIECELTDNVKMHDVIREMALWINSDFGKQQETICVKSGAHVRMIPNDIN
WEIVRTMSFTCTQIKKISCRSKCPNLSTLLILDNRLLVKISNRFFRFMPKLVVLDLSA
NLDLIKLPEEISNLGSLQYLNISLTGIKSLPVGLKKLRKLIYLNLEFTGVHGSLVGIA
ATLPNLQVLKFFYSCVYVDDILMKELQDLEHLKILTANVKDVTILERIQGDDRLASSI
RSLCLEDMSTPRVILSTIALGGLQQLAILMCNISEIRIDWESKERRELSPTEILPSTG
SPGFKQLSTVYINQLEGQRDLSWLLYAQNLKKLEVCWSPQIEEIINKEKGMNITKLHR
DIVVPFGNLEDLALRQMADLTEICWNYRTLPNLRKSYINDCPKLPEDIFVPLLPEKSP
SRFFFF"

CDS complement(join(24326. .24427,24512. .24622,24721. .24794,
24880. .24949,25034. .25081,25158. .25234,25321. .25447,
25525. .25617,25678. .25791,25883. .26128,26228. .26304,
26403. .26550,26640. .26724,27060. .27140,27291. .27395,
27483. .27637,27716. .27814,27937. .28009,28264. .28335,
28757. .29010,29183. .29278))
/note="emb|CAB86048.1
gene_id:MQD19.8
strong similarity to unknown protein"
/pseudo
/codon_start=1
/evidence=not_experimental

CDS complement(join(34991. .35221,35369. .35427,35507. .35569,
35676. .35739,35815. .36036))
/note="unnamed protein product; gene_id:MQD19.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB11302.1"
/db_xref="GI:10177943"
/translation="MATVTILSPKSIPKVTDSKFGARVSDQIVNVVKCGKSGRRLKLA
KLVSAAGLSQIEPDINEDPIGQFETNSIEMEDFKYGYYDGAHTYYEGEVQKGTFWGAI
ADDIAAVDQTNGFQGLISCMFLPAIALGMYFDAPGEYLFIGAALFTVVFCIIEMDKPD
QPHNFEPQIYKLERGARDKLINDYNTMSIWDFNDKYGDVWDFTIEKDDIATR"

CDS complement(37868. .38554)
/note="gene_id:MQD19.10"
/codon_start=1
/evidence=not_experimental
/product="non-LTR retrolelement reverse
transcriptase-like"
/protein_id="BAB11303.1"
/db_xref="GI:10177944"
/translation="MQGQLPISILWRIWKSRNKLVFQRKQNDWWRDIRDATNEAEEWI
TNGLCAQVPTPSTTNGTTRQFQSQWQKPHMGWIKCNYDGSFVNRVRGSTAAWIIRDDN
GVFKGAAQATGATVSSAFEAECQSLILTMQQLWIRGFRKVILEGDCKLLVNVLNGRGS
RFDAINWIHEIKVWRNKFDAIKFTWSHRSTNQPADILAKNQKHALSSYIVHSFIPLCI
SNALFSDYVS"

CDS join(40794. .41579,42574. .43377)
/note="gene_id:MQD19.11"
/codon_start=1
/evidence=not_experimental

Query Match 70.7%; Score 803; DB 15; Length 87286;
Best Local Similarity 99.4%; Pred. No. 1.5e-188;
Matches 806; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        310 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86386 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 86445

Qy        370 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86446 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 86505

Qy        430 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86506 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 86565

Qy        490 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86566 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 86625

Qy        550 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86626 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 86685

Qy        610 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86686 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 86745

Qy        670 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86746 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 86805

Qy        730 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86806 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 86865

Qy        790 GATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAG 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86866 GATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAG 86925

Qy        850 AGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTG 909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86926 AGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTG 86985

Qy        910 GAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTT 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86986 GAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTT 87045

Qy        970 GAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTT 1029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87046 GAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTT 87105

Qy       1030 CTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACT 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87106 CTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACT 87165

Qy       1090 CTTTAAGTCACCTCTTTGCTTACAAAAAAAA 1120
              ||||||||||||||||||||||  | |  ||
Db      87166 CTTTAAGTCACCTCTTTGCTTATGATATTAA 87196
```

```
RESULT 5
AF337545
LOCUS       AF337545               1104 bp    mRNA    linear   PLN 22-OCT-2002
DEFINITION  Brassica oleracea chlorophyllase 2 mRNA, complete cds.
ACCESSION   AF337545
VERSION     AF337545.1  GI:24210534
KEYWORDS    .
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 1104)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Molecular characterization and differential regulation of three
            chlorophyllase genes, members of the novel serine esterase gene
            family, in broccoli
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1104)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2001) Institute of Botany, Academia Sinica,
            Taipei, Taiwan 11529, Republic of China
FEATURES             Location/Qualifiers
     source          1. .1104
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /cultivar="Green king"
                     /db_xref="taxon:3712"
                     /tissue_type="floret"
     5'UTR           1. .41
     CDS             42. .1007
                     /EC_number="3.1.1.14"
                     /note="BoCLH2"
                     /codon_start=1
                     /product="chlorophyllase 2"
                     /protein_id="AAN51934.1"
                     /db_xref="GI:24210535"
                     /translation="MSSSSSRNAFVDGKYKPDLLTVDLASRCRCYKTTPSSSLTPPPP
                     PKSLLVATPVEEGEYPVVMLLHGYLLYNSFYSQLMLHVSSYGFIVIAPQLYNIAGPDT
                     IDEIKSTAEIIDWLSVGLNHFLPPQVTPNLSKFALTGHSRGGKTAFAVALKKFGYSSE
                     LKISAIIGVDPVDGTGKGKQTPPPVLTYEPNSFNLEKMPVLVIGSGLGELARNPLFPP
                     CAPTGVNHREFFQECQGPAWHFVAKDYGHLDMLDDDTKGLRGKSSYCLCKNGEERKPM
                     RRFIGGIVVSFLMAYLEDDDCELVKIKAGCHEGVPVEIQEFEVKK"
     3'UTR           1008. .1104
ORIGIN
```

```
  Query Match             65.6%;  Score 745;  DB 15;  Length 1104;
  Best Local Similarity   82.5%;  Pred. No. 4.9e-174;
  Matches  884;  Conservative    0;  Mismatches  170;  Indels   17;  Gaps    2;

Qy          2 AAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCAA 61
              | | ||  ||||  || ||||    | | |||| ||||||||||||| ||||||||||||
Db          1 AGAGAAAAAAAGTGAAAAAAATAACAAAGAGAAGAAAAAAAAATGTCATCTTCTTCATCAA 60

Qy         62 GAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCTC 121
              ||||||||||||  |||||||||||||||| || |||| |||||  |||| | | |||||||
Db         61 GAAACGCCTTTGTGGATGGCAAATACAAACCAGATCTTTTAACAGTGGATTTAGCATCTC 120

Qy        122 GTTGCTGC---------AAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGC 172
              | ||  ||         | |   || | ||||||    || | ||| |||| || |||
Db        121 GCTGTCGCTGCTACAAAACGACGCCATCTTCTTCTCTGACTCCGCCGCCTCCTCCCAAGT 180

Qy        173 AGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATG 232
              ||| |||||||| ||||||||||||||||||||| ||||||||||||||||||||||||
Db        181 CGCTTTTGGTGGCAACGCCGGTGGAGGAAGGAGAATATCCGGTGGTGATGCTCCTCCATG 240

Qy        233 GTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCT 292
              |||||||||||||||||| || || || |||||||||||||||||||||||  ||||||
Db        241 GTTACCTTCTCTACAACTCATTTTACTCCCAGCTTATGTTGCATGTCTCTTCCTATGGCT 300

Qy        293 TCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTA 352
              ||||  |||||||||| ||||||||||  || ||||||||||||| || || ||||| |
Db        301 TCATTGTCATCGCTCCGCAGTTATATAACATTGCCGGACCAGACACGATAGACGAGATAA 360

Qy        353 AATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGC 412
              |||||||||||||||||| |||||||||||| |||||||| |||||||||||| | |
Db        361 AATCAACGGCGGAGATTATTGATTGGTTATCAGTCGGACTTAACCACTTTCTTCCACCAC 420

Qy        413 AAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCG 472
              |||||||||| |||||||| ||||| || ||| ||||||||||||||| ||||||||||
Db        421 AAGTAACACCCAACCTATCTAAATTCGCACTCACCGGCCATAGCCGCGGCGGCAAAACCG 480

Qy        473 CGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGA 532
              | || || || ||||||||||||||||| ||||| ||| | |||||||||| || | |
Db        481 CATTCGCCGTGGCCTTAAAGAAATTTGGATACTCATCGGAACTAAAGATCTCAGCAATTA 540

Qy        533 TCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGG 592
              ||||| ||||||| || ||||| |||||||||| |||||||||||||||||||| ||
```

```
Db        541 TCGGTGTAGATCCGGTGGATGGGACAGGGAAAGGTAAACAAACCCCTCCTCCGGTTTTAA 600

Qy        593 CTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGC 652

Db        601 CCTATGAACCAAACTCATTTAACCTAGAAAAGATGCCTGTTCTAGTTATTGGTTCGGGAC 660

Qy        653 TTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACC 712

Db        661 TTGGTGAACTTGCCCGGAACCCATTGTTTCCACCGTGTGCACCTACGGGAGTGAACCACC 720

Qy        713 GAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGC 772

Db        721 GAGAGTTTTTCCAGGAATGTCAAGGTCCAGCGTGGCATTTTGTCGCAAAGGATTACGGAC 780

Qy        773 ATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGT 832

Db        781 ATTTGGACATGCTTGATGATGATACAAAAGGGCTTAGAGGGAAGAGTTCTTATTGTTTGT 840

Qy        833 GTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCAT 892

Db        841 GTAAGAATGGTGAAGAGAGAAAACCAATGAGGAGATTTATTGGTGGAATTGTTGTGTCGT 900

Qy        893 TTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTC 952

Db        901 TTTTGATGGCTTATCTGGAAGATGATGATTGTGAATTGGTGAAGATCAAAGCTGGGTGTC 960

Qy        953 ACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTC 1012

Db        961 ATGAAGGTGTTCCTGTTGAAATTCAAGAGTTTGAGGTTAAAAAGTAA--------TTTTC 1012

Qy       1013 TTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATG 1063

Db       1013 TTTAGAGCTTGTTTTTCCTATTTTCAATATACAATAGTCACACTGCTTATG 1063
```

```
RESULT 6
AF337546
LOCUS       AF337546                884 bp    mRNA    linear   PLN 22-OCT-2002
DEFINITION  Brassica oleracea chlorophyllase 3 mRNA, complete cds.
ACCESSION   AF337546
VERSION     AF337546.1  GI:24210536
KEYWORDS    .
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Molecular characterization and differential regulation of three
            chlorophyllase genes, members of the novel serine esterase gene
            family, in broccoli
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 884)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2001) Institute of Botany, Academia Sinica,
            Taipei, Taiwan 11529, Republic of China
FEATURES             Location/Qualifiers
     source          1. .884
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /cultivar="Green king"
                     /db_xref="taxon:3712"
                     /tissue_type="floret"
     5'UTR           1. .43
     CDS             44. .685
                     /EC_number="3.1.1.14"
                     /note="BoCLH3"
                     /codon_start=1
                     /product="chlorophyllase 3"
                     /protein_id="AAN51935.1"
                     /db_xref="GI:24210537"
                     /translation="MSPSFLFFTLFLIKEMSSSSSANSFEDGKYKTDLLTVGLSSCCW
                     KKPSSSPTPQSPPKRLLVATPVEEGEYPVVMLLHGYLLYNSFYSQLMLHVSSHGFIVI
                     APQLYSIAGPDTMDEIKSTAEIIDWLSVGLNHFLPPQVTPNLSKFALSGHSRGGKTAF
                     ALALKKFGYSSDLKISALIGIDVGTVFWTNGYGQYSGEFFEQFDCRNDRIVES"
     3'UTR           686. .884
ORIGIN
```

```
  Query Match           34.4%;  Score 390;  DB 15;  Length 884;
  Best Local Similarity   85.3%;  Pred. No. 8.4e-86;
  Matches  435;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy         34 ACAAAAAAAAATGTCCTCTTCTTCATCAAGAAACGCCTTTGAAGATGGCAAATACAAATCA 93

Db         80 ATAAAGGAAATGTCCTCTTCATCATCAGCAAACTCCTTTGAGGACGGCAAATACAAAACA 139

Qy         94 AATCTCTTAACCTTGGACTCATCATCTCGTTGCTGCAAAATAACACCGTCTTCTAGAGCT 153

Db        140 GATCTTTTAACAGTAGGCTTATCATCTTGCTGCTGGAAAAAGCCCTCCTCTTCTCCGACT 199

Qy        154 TCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCG 213

Db        200 CCGCAGTCTCCGCCGAAGAGGCTTTTGGTGGCAACGCCGGTGGAGGAAGGAGAATATCCG 259

Qy        214 GTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTG 273

Db        260 GTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCATTTTATTCCCAGCTTATGTTG 319

Qy        274 CATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCA 333

Db        320 CATGTCTCTTCCCATGGCTTCATTGTCATCGCTCCGCAGTTATATAGCATTGCCGGACCA 379

Qy        334 GACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTT 393

Db        380 GACACCATGGATGAGATAAAATCAACGGCAGAGATTATTGATTGGTTATCGGTCGGACTA 439

Qy        394 AATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCAT 453

Db        440 AACCACTTTCTTCCACCACAAGTAACACCAAACCTATCCAAGTTCGCACTCTCCGGCCAT 499

Qy        454 AGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAAT 513

Db        500 AGCCGTGGTGGGAAGACCGCATTTGCCTTGGCCTTAAAGAAATTTGGATACTCGTCCGAC 559

Qy        514 CTAAAGATCTCGACATTGATCGGTATAGAT 543

Db        560 CTAAAGATCTCGGCATTGATAGGTATAGAT 589
```

```
RESULT 7
AX411603
LOCUS       AX411603                1216 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 3 from Patent WO0229022.
ACCESSION   AX411603
VERSION     AX411603.1  GI:21444160
KEYWORDS    .
SOURCE      Vitis sp.
  ORGANISM  Vitis sp.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; Vitaceae; Vitis.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 3 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1216
                     /organism="Vitis sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3604"
ORIGIN

  Query Match           32.4%;  Score 367.4;  DB 6;  Length 1216;
```

```
  Best Local Similarity   63.9%;  Pred. No. 3.4e-80;
  Matches  608;  Conservative    0;  Mismatches  331;  Indels   12;  Gaps    3;

Qy         50 CTTCTTCATCAAGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGG 109
Db          1 CTTCTACATGTTCCACCACTGCTGCAAATGTTTTGAGATTGGAAAGCACATCACAGTGC 60

Qy        110 ACTCATCATCTCGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAA 169
Db         61 TTCTAAGAGCTGAACCAGGCACTTGCACCACCAAGTCATCTCTTCCTGTTCCCCCTCCAC 120

Qy        170 AGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCC 229
Db        121 TGCAACTCTTGATTGCTACGCCCTCTGAAGCAGGGGAGTTCCCGCTGCTGCTCCTCCTTC 180

Qy        230 ATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATG 289
Db        181 ATGGTTATCTTCTCTATAACTCTTTCTACTCCCAGCTCATCCAACACATAGCCTCTCATG 240

Qy        290 GCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGA 349
Db        241 GTTTCATTGTTCTTGCTCCTCAGTTATACACTGTGGCTGGACCAGATTCAAGCGAAGAGA 300

Qy        350 TTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAG 409
Db        301 TCAAGTCCGCAGCTGCTTTAACAAATTGGTTATCCAAAGGACTCCATGACTTACTTCCTC 360

Qy        410 CGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAA 469
Db        361 CCCATGTTCGGCCAAATTTAAGCAAACTAGGACTTGCCGGCCATAGTCGTGGAGGCAAAA 420

Qy        470 CCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACAT 529
Db        421 CTGCTTTTGCTCTAGCACTGAGAAAA------GCATCCACTTCTCTGAAATTTTCAGCCT 474

Qy        530 TGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGT 589
Db        475 TGATAGGCATAGACCCGGTGGATGGAATGGACAAAGGGAAACAAACCCCTCCACCGGTAC 534

Qy        590 TGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGG 649
Db        535 TCACCTATGTTCCTCATTCATTTGATCTAGAC---ATGGCAGTGATGGTAATTGGTTCGG 591

Qy        650 GGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATC 709
Db        592 GTTTGGGTGAAGTGAAAAGGAACCCTCTGTTCCCTCCTTGTGCCCCCAAGGGCGTAAACC 651

Qy        710 ACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATG 769
Db        652 ATGAGGACTTCTTTAAAGAATGCCGTGAACCAGCTTGTTATTTTCTTGCCAAGGACTATG 711

Qy        770 GGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTT 829
Db        712 GCCATCTTGACATGCTAGACGATGAGACTAATGGAATTAGAGGGAAAGCTACACATTGTT 771

Qy        830 TGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTAT 889
Db        772 TGTGTAAAAAATGGGAAGTCTAGAGAACCCATGAGGAGGTTTGTTGGAGGCATTGTGATTG 831

Qy        890 CATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGT 949
Db        832 CATTTATGAAAGCTTATTTGGAAGGCGATAACAGCAGTCTAATCTCCATTAGAGATGGG- 890

Qy        950 GTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAA 1000
Db        891 --CATGCTACTGCACCAGTGGAGCTTCAAAATGTTGAGTTTCTCGTGTGAA 939


RESULT 8
BT013523
LOCUS       BT013523              1159 bp    mRNA    linear   PLN 11-MAY-2004
DEFINITION  Lycopersicon esculentum clone 132241R, mRNA sequence.
ACCESSION   BT013523
VERSION     BT013523.1  GI:47104938
KEYWORDS    FLI_CDNA.
SOURCE      Lycopersicon esculentum (Solanum lycopersicum)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 1159)
  AUTHORS   Kirkness,E.F., Wang,W. and Vazeille,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..1159
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
                     /clone="132241R"
                     /tissue_type="fruit"
                     /note="TZXBH55"
ORIGIN

  Query Match             29.8%;  Score 338.6;  DB 15;  Length 1159;
  Best Local Similarity   64.0%;  Pred. No. 5e-73;
  Matches  547;  Conservative    0;  Mismatches  299;  Indels    9;  Gaps    2;

Qy        160 TCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTG 219
Db        149 TCTCCTCCAAAATCTCTGTTAATTGGGACCCCATCAGAAGCAGGAAACTTCCCAGTATTG 208

Qy        220 ATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTC 279
Db        209 ATATTTCTCCATGGCTATCTTCTTTACAATTCTTTCTATTCACAACTTATTCAACATCTC 268

Qy        280 TCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACA 339
Db        269 TCTTCTCATGGCTTCATTGTTGTTGCTCCTCAGTTATATTTAGTGGAAGGAGCAGATGCA 328

Qy        340 ATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCAC 399
Db        329 ACCGAGGATATCAAATCGACAGCTGAAGTCACCCATTGGTTATCTGAAGGATTGCAGCAT 388

Qy        400 TTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGC 459
Db        389 CACCTTCCACCTGATGTTGAGCCAAACTTGACGAAACTTGGACTAGCAGGCCATAGCCGT 448

Qy        460 GGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAG 519
Db        449 GGGGGAAAAGTTGCATTTTCGCTAGCTCTTGGTAGACTTG------CTAGTGACCTGAAA 502

Qy        520 ATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCT 579
Db        503 TTTTCAGCGTTGATCGGTGTTGATCCTGTGGATGGAATGGAGAAAGGAAAGCAAACTCCG 562

Qy        580 CCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTG 639
Db        563 CCATCAGTTCTCACCTACGTTCCGCGTTCCTTCATTAATCTTGATATGCCTGTGATGGTA 622

Qy        640 ATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCC 699
Db        623 ATTGGCTCGGGTTTGGGAGAAGTTAAAAAGAATCCTCTGTTCCCTGCTTGTGCTCCTAAA 682

Qy        700 GGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCG 759
Db        683 GGGGTGAATCATCGCGATTTCTACAACGAATGTTGCAAGCCAGCTTGTTACTTTGTTGCA 742

Qy        760 AAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGT 819
Db        743 AAGGATTATGGACATAATGATATGTTAGATGAGGAGACCAAAGGGATTCGAGGCAAGGCT 802

Qy        820 TCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGA 879
Db        803 ACTTATTGTTTGTGCAAGAAAGGGAAATCTAGAGAGCCCATGAGGAGATTTGTTGGAGGG 862
```

```
Qy        880 CTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATC 939
              || | ||  |||||||| | ||||||||||||||  ||    |  | ||  |    ||
Db        863 GTTTTGGTCGCATTTTTGGAAGCTTATTTGGAAGGTAATTCCAGCCACTTGATAGCTATT 922

Qy        940 AAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAA 999
              | |||||   | || |  |   | || || ||  | |||||  ||||  || |  | |||
Db        923 AGAGATG---GCCATGTTGCACTGCCTGTAGAGCTCCAAGATATTGATTTTCTTGTCTAA 979

Qy       1000 ACATAAGTTTTTCTT 1014
              |    ||| |  ||
Db        980 GCTATGGTTGTGGTT 994


RESULT 9
AP007972
LOCUS       AP007972              128855 bp    DNA     linear   HTG 28-DEC-2004
DEFINITION  Lotus corniculatus var. japonicus chromosome 5 clone LjT44A18, ***
            SEQUENCING IN PROGRESS ***, 45 unordered pieces.
ACCESSION   AP007972
VERSION     AP007972.1  GI:56806280
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Kaneko,T., Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. XI. Sequence
            Features and Mapping of Nine hundred twenty-one TAC Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 128855)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
            Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
            Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
            Fax:81-438-52-3934)
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 45 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1715: contig of 1715 bp in length
            *     1716     1815: gap of unknown length
            *     1816     2641: contig of 826 bp in length
            *     2642     2741: gap of unknown length
            *     2742     3555: contig of 814 bp in length
            *     3556     3655: gap of unknown length
            *     3656     4548: contig of 893 bp in length
            *     4549     4648: gap of unknown length
            *     4649     6180: contig of 1532 bp in length
            *     6181     6280: gap of unknown length
            *     6281     7099: contig of 819 bp in length
            *     7100     7199: gap of unknown length
            *     7200     8339: contig of 1140 bp in length
            *     8340     8439: gap of unknown length
            *     8440     9486: contig of 1047 bp in length
            *     9487     9586: gap of unknown length
            *     9587    11121: contig of 1535 bp in length
            *    11122    11221: gap of unknown length
            *    11222    12794: contig of 1573 bp in length
            *    12795    12894: gap of unknown length
            *    12895    14686: contig of 1792 bp in length
            *    14687    14786: gap of unknown length
            *    14787    16382: contig of 1596 bp in length
            *    16383    16482: gap of unknown length
            *    16483    17841: contig of 1359 bp in length
            *    17842    17941: gap of unknown length
            *    17942    19137: contig of 1196 bp in length
            *    19138    19237: gap of unknown length
            *    19238    21270: contig of 2033 bp in length
            *    21271    21370: gap of unknown length
            *    21371    22827: contig of 1457 bp in length
            *    22828    22927: gap of unknown length
            *    22928    24839: contig of 1912 bp in length
            *    24840    24939: gap of unknown length
            *    24940    27581: contig of 2642 bp in length
            *    27582    27681: gap of unknown length
            *    27682    29435: contig of 1754 bp in length
            *    29436    29535: gap of unknown length
            *    29536    31538: contig of 2003 bp in length
            *    31539    31638: gap of unknown length
            *    31639    33866: contig of 2228 bp in length
            *    33867    33966: gap of unknown length
            *    33967    36055: contig of 2089 bp in length
            *    36056    36155: gap of unknown length
            *    36156    38250: contig of 2095 bp in length
            *    38251    38350: gap of unknown length
            *    38351    40933: contig of 2583 bp in length
            *    40934    41033: gap of unknown length
            *    41034    43319: contig of 2286 bp in length
            *    43320    43419: gap of unknown length
            *    43420    45644: contig of 2225 bp in length
            *    45645    45744: gap of unknown length
            *    45745    47854: contig of 2110 bp in length
            *    47855    47954: gap of unknown length
            *    47955    50383: contig of 2429 bp in length
            *    50384    50483: gap of unknown length
            *    50484    53820: contig of 3337 bp in length
            *    53821    53920: gap of unknown length
            *    53921    56225: contig of 2305 bp in length
            *    56226    56325: gap of unknown length
            *    56326    59890: contig of 3565 bp in length
            *    59891    59990: gap of unknown length
            *    59991    62614: contig of 2624 bp in length
            *    62615    62714: gap of unknown length
            *    62715    65635: contig of 2921 bp in length
            *    65636    65735: gap of unknown length
            *    65736    69840: contig of 4105 bp in length
            *    69841    69940: gap of unknown length
            *    69941    73812: contig of 3872 bp in length
            *    73813    73912: gap of unknown length
            *    73913    79015: contig of 5103 bp in length
            *    79016    79115: gap of unknown length
            *    79116    83508: contig of 4393 bp in length
            *    83509    83608: gap of unknown length
            *    83609    87605: contig of 3997 bp in length
            *    87606    87705: gap of unknown length
            *    87706    92715: contig of 5010 bp in length
            *    92716    92815: gap of unknown length
            *    92816    98011: contig of 5196 bp in length
            *    98012    98111: gap of unknown length
            *    98112   102876: contig of 4765 bp in length
            *   102877   102976: gap of unknown length
            *   102977   108287: contig of 5311 bp in length
            *   108288   108387: gap of unknown length
            *   108388   112790: contig of 4403 bp in length
            *   112791   112890: gap of unknown length
            *   112891   127873: contig of 14983 bp in length
            *   127874   127973: gap of unknown length
            *   127974   128855: contig of 882 bp in length.
FEATURES             Location/Qualifiers
     source          1. .128855
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /chromosome="5"
                     /clone="LjT44A18"
```

```
                /clone_lib="LjT library"
                /note="TAC clone:TM1301, synonym:Lotus japonicus"
     gap        1716. .1815
                /estimated_length=unknown
     gap        2642. .2741
                /estimated_length=unknown
     gap        3556. .3655
                /estimated_length=unknown
     gap        4549. .4648
                /estimated_length=unknown
     gap        6181. .6280
                /estimated_length=unknown
     gap        7100. .7199
                /estimated_length=unknown
     gap        8340. .8439
                /estimated_length=unknown
     gap        9487. .9586
                /estimated_length=unknown
     gap        11122. .11221
                /estimated_length=unknown
     gap        12795. .12894
                /estimated_length=unknown
     gap        14687. .14786
                /estimated_length=unknown
     gap        16383. .16482
                /estimated_length=unknown
     gap        17842. .17941
                /estimated_length=unknown
     gap        19138. .19237
                /estimated_length=unknown
     gap        21271. .21370
                /estimated_length=unknown
     gap        22828. .22927
                /estimated_length=unknown
     gap        24840. .24939
                /estimated_length=unknown
     gap        27582. .27681
                /estimated_length=unknown
     gap        29436. .29535
                /estimated_length=unknown
     gap        31539. .31638
                /estimated_length=unknown
     gap        33867. .33966
                /estimated_length=unknown
     gap        36056. .36155
                /estimated_length=unknown
     gap        38251. .38350
                /estimated_length=unknown
     gap        40934. .41033
                /estimated_length=unknown
     gap        43320. .43419
                /estimated_length=unknown
     gap        45645. .45744
                /estimated_length=unknown
     gap        47855. .47954
                /estimated_length=unknown
     gap        50384. .50483
                /estimated_length=unknown
     gap        53821. .53920
                /estimated_length=unknown
     gap        56226. .56325
                /estimated_length=unknown
     gap        59891. .59990
                /estimated_length=unknown
     gap        62615. .62714
                /estimated_length=unknown
     gap        65636. .65735
                /estimated_length=unknown
     gap        69841. .69940
                /estimated_length=unknown

Query Match             24.8%;  Score 281.6;  DB 14;  Length 128855;
Best Local Similarity   66.1%;  Pred. No. 5.8e-59;
```

```
  Matches  423;  Conservative    0;  Mismatches  214;  Indels    3;  Gaps    1;

Qy        305 CTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGG 364
              ||   ||||| |||     | || ||||| ||    ||  |  |||||| | || || || |
Db      79988 CTTTACAGTTGTATGCAGTGGCTGGACCTGATGTAAGTGGAGAGATTCATTCTACAGCTG 80047

Qy        365 AGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAA 424
              || |   ||||| |||| | |||||| |    | || ||||||  |  | ||||||||||
Db      80048 CAATAACAAATTGGCTATCTGAAGGACTCAGCAAGTTCCTTCCACCAAATGTAACACCAA 80107

Qy        425 ACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCG 484
              |  |   ||| || || ||| ||||||| || ||||| ||||||||  || ||||| || |
Db      80108 ATTTCAGCAAGTTAGCACTCGCCGGCCACAGTCGCGGAGGCAAAACAGCTTTTGCTGTTG 80167

Qy        485 CCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATC 544
              |  | |  ||| |       | |  |  |||| ||| | ||  ||||  | ||  | || |
Db      80168 CTCTGAGAAAACTAAACATCACAACTGATCTCAAGTTTTCAGCATTAGTAGGAGTGGACC 80227

Qy        545 CAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAA 604
              |||| ||||||    |  | ||| || |||||||| || || ||  |  | ||  ||||
Db      80228 CAGTTGATGGACTGGACAGAGGAAAGCAAACCCCCCCCACCAGTTCTCACCTATGTTCCTC 80287

Qy        605 ACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCG 664
              | ||||||||  | |||   | |||   | | || || ||||| || || || ||
Db      80288 ATTCATTTGATTTCGAC---ATGCCAGCAATGGTTATAGGTTCAGGTCTAGGGGACGTGA 80344

Qy        665 CTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTC 724
              ||||| || ||||| || || ||||| ||        || || ||     || |||||
Db      80345 AAAGGAATCCTTTATTTCCCCCTTGTGCTCCCAAAACTGTCAACCATGAGGATTTCTTCA 80404

Qy        725 GGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGC 784
              |||||| |    || || ||| |||| || || |||||||||| ||| | |||||||
Db      80405 ATGAATGTAACAAACCTGCTTGGTATTTTGTGGCTAAGGATTATGGACATGTTGACATGC 80464

Qy        785 TTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTG 844
              | |||||||||| || || || | ||||||   | | |||||| | || ||||||| |
Db      80465 TAGATGATGATACCAATGGAATAATAGGGAAAGCTACCTATTGTCTATGCAAGAATGGAG 80524

Qy        845 AAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTT 904
              |    |||| |||||||||| ||| ||||| |||||||| ||  |||| ||||| ||||
Db      80525 AGTCAAGGAAACCAATGAGGACATTTGTTGGAGGACTTGTAGTTGCATTCTTGAAAGCTT 80584

Qy        905 ATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGA 944
              |  || | ||  ||| |  | || ||| |    || |||||
Db      80585 ACCTGCAGGGTGATAACAGGGACTTATTGGCCATAAAAGA 80624
```

```
RESULT 10
AX411613
LOCUS       AX411613               1444 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 13 from Patent WO0229022.
ACCESSION   AX411613
VERSION     AX411613.1  GI:21444165
KEYWORDS    .
SOURCE      Tulipa gesneriana
  ORGANISM  Tulipa gesneriana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
            Tulipa.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 13 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1444
                     /organism="Tulipa gesneriana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:13306"
ORIGIN

  Query Match             22.6%;  Score 256.8;  DB 6;  Length 1444;
  Best Local Similarity   61.1%;  Pred. No. 1e-52;
```

```
  Matches  474;  Conservative    0;  Mismatches  287;  Indels   15;  Gaps    3;

Qy        154 TCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCG 213
              ||  |  | || |  ||||  |   || | || || ||  | ||||||||| || || ||
Db        199 TCCGCACCACCTCGGAAGCCACCACTGATCGCAACTCCATTTGAGGAAGGCGAGTACCCA 258

Qy        214 GTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTG 273
              | |  | ||||| ||||| | | |  ||  ||||| | ||||| ||| |||||||  |
Db        259 ACGCTTCTACTCCTTCATGGATTCATGCTTCACAACACTTTCTACTCTGAGCTTATCCAG 318

Qy        274 CATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCA 333
              ||  || | || |||||||| ||  |  | |  ||||||||||||  | | ||   |
Db        319 CACATCGCATCCCATGGCTTTATTGTTGTTGTACCTCAGTTATATCTTGTAGCTACATGT 378

Qy        334 GACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTT 393
              ||  |  | | |||  || || ||  | ||  | |  |  |||||||||     | |||||
Db        379 GATAGTACGAATGGCATCAAGTCTGCTGCAAAAACAACAGATTGGTTGAAGGATGGACTG 438

Qy        394 AATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCAT 453
              |  |  |||| ||| |  |||| | |||| |||||   ||| | |  ||   ||| |||
Db        439 CAAGATGTTCTCCCAACAAAAGTCAGACCAGACCTAAAGAAACTCGGACTGAGCGGACAT 498

Qy        454 AGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAAT 513
              ||||| || ||||||   || |||||  | ||  ||    |       |     | ||
Db        499 AGCCGTGGCGGCAAAGATGCATTTGCTCTTGCACTAGGATATGCGAAG------ACTACA 552

Qy        514 CTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAA 573
              |||   ||||  |  ||||||| || ||||| || || |||  | | || ||  | |||
Db        553 TTAAGCTTCTCAGCGCTGATCGGGATCGATCCTGTTGACGGAGTACGAAAGGGACACCAA 612

Qy        574 ACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATA 633
              |||  |||||| ||  ||  |||  | ||  |||| || || ||  ||| | |||   |
Db        613 ACCAATCCTCCTGTACTGAATTATATCCCTCACTCTTTGGAACT---CAAGATGCCGTCA 669

Qy        634 CTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCA 693
              | ||  |||||  | ||  | ||||||      ||||||  |||   |      |||||
Db        670 TTAGTTATCGGAACAGGTTTAGGTGAATTGAAGAGGAACCTATTTGCC------TGTGCC 723

Qy        694 CCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTC 753
              |||   ||||| |||||||  |||  || |      || |||       ||  | ||  |||||
Db        724 CCTAAGGGAGTAAATCACCAAGATTTTTACGATGAGTGTTCTTCCCCGTCTTGCCATTTT 783

Qy        754 GTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGG 813
              ||||| ||||| |||| ||  |||||||||| ||||||  |||| | ||  |  ||||||
Db        784 GTTGCTAAGGACTATGGTCACATGGACATGCTAGATGATCATACTGACGGACCTAGAGGT 843

Qy        814 AAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTT 873
              |     || || ||  |||||  ||||| |||   || | ||| || |    || ||
Db        844 ATTGTGTCATACTGCATGTGTCGAAATGGAGAATCAAGAACACCCATGCGACTGTTTGTA 903

Qy        874 GGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATT 929
              |||||  | || ||  | ||| |||| || |  ||  |||| ||| | |  || ||
Db        904 GGTGGGATCGTGGTTGCTTTTCTGAAAGCCTGCTTACAAGGTGATAAACAGGATTT 959


RESULT 11
AY292526
LOCUS       AY292526               1157 bp    mRNA    linear   PLN 05-JUN-2003
DEFINITION  Ginkgo biloba chlorophyllase (CLH) mRNA, complete cds.
ACCESSION   AY292526
VERSION     AY292526.1  GI:31415499
KEYWORDS    .
SOURCE      Ginkgo biloba (maidenhair tree)
  ORGANISM  Ginkgo biloba
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Ginkgophyta; Ginkgoales; Ginkgoaceae; Ginkgo.
REFERENCE   1  (bases 1 to 1157)
  AUTHORS   Okazawa,A., Tang,L., Fukusaki,E.-I. and Kobayashi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2003) Department of Biotechnology, Osaka
            University, 2-1, Yamadaoka, Suita, Osaka 565-0871, Japan
FEATURES             Location/Qualifiers
     source          1. .1157
                     /organism="Ginkgo biloba"
                     /mol_type="mRNA"
                     /db_xref="taxon:3311"
     gene            1. .1157
                     /gene="CLH"
                     /note="GbCLH"
     CDS             55. .1083
                     /gene="CLH"
                     /note="GbCLH"
                     /codon_start=1
                     /product="chlorophyllase"
                     /protein_id="AAP44978.1"
                     /db_xref="GI:31415500"
                     /translation="MVLVKDVFSEGPLPVQILAIPQANSSPCSKLADKNGTATTPSPC
                     RPPKPLLIALPSQHGDYPLILFFHGYVLLNSFYSQLLRHVASHGYIAIAPQMYSVIGP
                     NTTPEIADAAAITDWLRDGLSDNLPQALNNHVRPNFEKFVLAGHSRGGKVAFALALGR
                     VSQPSLKYSALVGLDPVDGMGKDQQTSHPILSYREHSFDLGMPTLVVGSGLGPCKRNP
                     LFPPCAPQGVNHHDFFYECVAPAYHFVASDYGHLDFLDDDTKGIRGKATYCLCKNGEA
                     REPMRKFSGGIVVAFLQAFLGDNRGALNDIMVYPSHAPVKIEPPESLVTEDVKSPEVE
                     LLRRAVCR"
ORIGIN

  Query Match             21.8%;  Score 247.4;  DB 15;  Length 1157;
  Best Local Similarity   59.9%;  Pred. No. 2.3e-50;
  Matches  476;  Conservative    0;  Mismatches  301;  Indels   18;  Gaps    3;

Qy        163 CCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATG 222
              || || || |  ||| || | |||   ||      |  | ||||||||||| |  |  |
Db        190 CCTCCTAAACCCCTGCTGATCGCTCTTCCTTCCCAACATGGAGATTATCCTCTCATCCTC 249

Qy        223 CTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCT 282
              |  |||| || ||  | |||  ||| |||||||||||||| ||  ||   |||||  ||
Db        250 TTTTTCCACGGCTATGTACTCCTCAATTCCTTCTATTCTCAACTCTTGCGCCATGTTGCT 309

Qy        283 TCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATG 342
              || ||||| | |||| ||| |||||||||| | || ||| |    || ||| | ||| |
Db        310 TCCCATGGATACATCGCCATAGCTCCTCAGATGTACAGTGTAATTGGCCCAAATACGACT 369

Qy        343 GATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTT 402
              || ||         | ||||  ||||  || ||||||  ||  |||||      |   |
Db        370 CCAGAAATAGCCGATGCAGCGGCCATTACAGACTGGTTACGAGATGGACTCTCGGATAAT 429

Qy        403 CTTCC------------AGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450
              |||||             || || |  || ||  |    |||||||  ||  | ||
Db        430 CTTCCGCAAGCTTTAAACAATCATGTGAGGCCCAATTTTGAGAAATTTGTGCTAGCGGGG 489

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCG 510
              ||    ||||| || |||   || |||||  | ||| ||      | |  | |  | ||
Db        490 CACTCGCGCGGGGGTAAAGTGGCATTTGCACTTGCCCTAGGTCGAGTCTCGCAGCCATC- 548

Qy        511 AATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAA 570
              | |||||  |||| |  |  | ||| | ||||||||||||||||  || ||||   ||
Db        549 --TTTAAAGTACTCGGCCCTTGTAGGTCTTGATCCAGTCGATGGAATGGGAAAAGATCAA 606

Qy        571 CAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCT 630
              ||||||  || |||  |  || | |||       | || |||||  | |   | |||
Db        607 CAAACCAGTCATCCTATTCTGTCATACAGAGAGCATTCCTTTGATTTGG---GTATGCCA 663

Qy        631 ATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGT 690
              | | | ||| | ||||||||  || |||      |    | |||||  | ||||| || |||
Db        664 ACATTAGTGGTAGGTTCGGGCCTGGGTCCGTGCAAAAGAAACCCTCTCTTCCCTCCCTGT 723

Qy        691 GCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCAT 750
              || || |  || ||  || ||||  || |||||     ||||||    | ||| ||  | |||
Db        724 GCTCCCCAAGGTGTTAACCACCATGATTTCTTCTACGAATGTGTCGCTCCTGCCTATCAT 783

Qy        751 TTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGA 810
              || |||||   |||||||||||| | ||  |  | || || || || || ||||| || |||
Db        784 TTTGTTGCCTCTGATTATGGGCATCTTGATTTCTTAGACGACGACACCAAAGGAATAAGA 843

Qy        811 GGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTC 870
              || |||  | |||||||  | ||||||||||| |||| ||    ||||||| |||  ||
Db        844 GGAAAGGCTACTTATTGCCTCTGTAAGAATGGGGAAGCAAGAGAGCCAATGCGGAAGTTT 903
```

```
Qy        871 GTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTA 930
              ||||||  ||||  ||  |||||  |  | ||  |  |  |  |   ||  |  |  |  |
Db        904 AGCGGTGGAATTGTGGTTGCATTTCTTCAAGCATTTCTTGGTGATAATCGTGGAGCCCTG 963

Qy        931 GTTAAGATCAAAGAT 945
              |  |  ||  |  | |
Db        964 AATGATATTATGGTT 978


RESULT 12
AX411601
LOCUS       AX411601                 987 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 1 from Patent WO0229022.
ACCESSION   AX411601
VERSION     AX411601.1  GI:21444159
KEYWORDS    .
SOURCE      Vitis sp.
  ORGANISM  Vitis sp.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; Vitaceae; Vitis.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 1 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .987
                     /organism="Vitis sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3604"
ORIGIN

  Query Match             20.9%;  Score 237.4;  DB 6;  Length 987;
  Best Local Similarity   56.3%;  Pred. No. 7.1e-48;
  Matches  551;  Conservative    0;  Mismatches  411;  Indels   17;  Gaps    5;

Qy        158 CGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGG 217
              | ||  ||  ||  ||||   |||||  |  |  |||  ||       ||||    ||  ||  ||
Db         23 CTTCCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTC 82

Qy        218 TGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATG 277
              |  ||  |  ||  |||||  |  |    |||  ||||  |||||||  |||||||  |     |
Db         83 TCTTGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCA 142

Qy        278 TCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACA 337
              |  |||||  |||||  |||||   |   |  ||||||||||||||   |  |        ||
Db        143 TTTCTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTG 202

Qy        338 CAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATC 397
              |||   |  |||||  ||||||  |  ||  |    |  |    |||||  ||||   |||  |||  |
Db        203 GAATTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAAT 262

Qy        398 ACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCC 457
                 |  ||  ||||    |  ||  |  ||||  |||||   |||   ||||  ||  ||  |||||  |||
Db        263 CTGTGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCA 322

Qy        458 GCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAA 517
              |  ||  ||  ||  ||  ||  |||||  |  ||  ||  |  |    ||      |||       ||  |
Db        323 GAGGGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGCTGATACATCC------CTCA 376

Qy        518 AGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCC 577
              |  ||||  |  |  |  ||  ||||  ||  ||  |  |||      | |||  |    |||||
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        578 CTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTG 637
               |||      |  |  ||  ||||  |  ||  ||  |  ||||    ||  ||  |    |
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAG---CAATCCCAGTTTGCG 493

Qy        638 TGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTC 697
              |  |||||  |||||  |  ||  ||      |  |||||    ||       ||  |||||  ||
Db        494 TAATCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAG 553
```

```
Qy        698 CCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTG 757
              ||||||||  ||    ||||||  ||  |  ||  |||  ||   |||    |   ||  ||  ||
Db        554 ATGGAGTGAACCATGTAGAGTTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGA 613

Qy        758 CGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGA 817
              |  |  ||  |||||  ||  |||||||||  |  ||||||  ||      |||     |  |||  ||
Db        614 CTACTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGA 673

Qy        818 GTTCTTATTGTTTGTGTAAGAATGGTGAAG---AGAGGAGACCAATGAGGAGATTCGTTG 874
               |||    ||  |  |  ||  ||||  |||   |  |      |||    ||  ||||||||||   ||  |
Db        674 TTTCGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGG 733

Qy        875 GTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTA 934
              ||||  |||  ||||   ||||  ||||||||||||||||||||   |   |  |  ||  |    |  |
Db        734 GTGGCCTTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATT---TCA 790

Qy        935 AGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCA 994
              |   ||  |  ||      |  ||    |  |||  ||  |  ||  |      |||||  |  ||
Db        791 AAGCCATTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCTGTTGAGTTCATAG 850

Qy        995 TGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTG 1054
              |  |  |  |    |  |  |    ||  |  |    | |||  |         |  ||||
Db        851 AAGCATGAAATCATGGATTCTTTGTAATGCTATAAGGAGTGTGA--CAGCAGATTTTTTT 908

Qy       1055 TTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAA 1114
              ||  |  |  |||   |||  | |||  |  |  |  |        |||||  | ||
Db        909 TTCGGTTCGTCTTTGGGAACAAAGATTCAAATAAACGAAACTGTGTTGTATTGCTAAAAA 968

Qy       1115 AAAAAAAAAAAAAAAAAAA 1133
              ||||||||||||||||||
Db        969 AAAAAAAAAAAAAAAAAAA 987


RESULT 13
ATH524543
LOCUS       ATH524543                242 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            076H05.
ACCESSION   AJ524543
VERSION     AJ524543.1  GI:26792779
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
   PUBMED   12446565
REFERENCE   2  (bases 1 to 242)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .242
```

```
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="076H05"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Wassilewskija"
     misc_feature    1. .242
                     /note="T-DNA flanking sequence
                     left border"
ORIGIN

  Query Match             19.5%;  Score 221.6;  DB 15;  Length 242;
  Best Local Similarity   98.2%;  Pred. No. 6.4e-44;
  Matches  224;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         87 CAAATCAAATCTCTTAACCTTGGACTCATCATCTCGTTGCTGCAAAATAACACCGTCTTC 146
Db          1 CAAATCAAATCTCTTAACCTTGGACTCATCATCTCGTTGCTGCAAAATAACACCGTCTTC 60

Qy        147 TAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGA 206
Db         61 TCGGGCTTCACCGTCTCCGCCGAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGA 120

Qy        207 TTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCT 266
Db        121 TTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCT 180

Qy        267 TATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTT 314
Db        181 TATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGGT 228


RESULT 14
AX411605
LOCUS       AX411605                1302 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 5 from Patent WO0229022.
ACCESSION   AX411605
VERSION     AX411605.1  GI:21444161
KEYWORDS    .
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Thorpe,C., Cahoon,E.B. and Cahoon,R.E.
  TITLE     Chlorophyllases
  JOURNAL   Patent: WO 0229022-A 5 11-APR-2002;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .1302
                     /organism="Zea mays"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4577"
ORIGIN

  Query Match             17.8%;  Score 201.6;  DB 6;  Length 1302;
  Best Local Similarity   54.6%;  Pred. No. 5.5e-39;
  Matches  449;  Conservative    0;  Mismatches  364;  Indels    9;  Gaps    2;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
Db        241 CGTCCAGCGCGGCCAGCCTGCCGCCGAAGCCGCTGCTCGTCGCCGCGCCGCGCGAGACTG 300

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
Db        301 GGGAGTACCCGGTGATCCTGTTCCTACACGGCTACCTCGCCGTCAACTCCTTCTACTCCC 360

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
Db        361 AGCTGTTCGAGCACGTCGCCTCCCATGGCTTTATCGTCGTCGGACCTCAGCTGTACACCA 420

Qy        323 TCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTAT 382
Db        421 TATCTGGGGCCGACACCACCGAGGAGATCAACTCAGCGGCGGCCGTCATCGACTGGCTAG 480

Qy        383 CAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCC 442
Db        481 CCACCGGGCTGCCGTCAACTCTGCCACTCGGCGTCCGCGCGAACCTAACCAAGGTGTCCA 540

Qy        443 TCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGT 502
Db        541 TCTCCGGCCACAGTCGCGGCGGGAAGGTGGCGTTCGCGCTGGCGTTGGGCCACGCCAAGG 600

Qy        503 ACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGA 562
Db        601 CCAAGCTCGCTGTCCCTCTCGCCGCCGTCGTCGCCGTGGACCCGGTGGACGGCATGGGCG 660

Qy        563 AAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTGACCTAGACA 622
Db        661 TGGGCAAGCAGACACCGCCGCCGATCCTCACGGGCAGGCACGGCTCGCTGCACGTGG--- 717

Qy        623 AAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCC 682
Db        718 GTGCCCCCGCCATGGTCATCGGCACGGGGCTCGGCGAGCTGCCCCGCGGCTCACTGCTCC 777

Qy        683 CACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCT------TTCGGGAATGTCAAG 736
Db        778 CGCCGTGCGCGCCCCGGGGCGTCAGTCACGCGGCCTTCTACGACGAGCTGGACGGCGCGG 837

Qy        737 GTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATA 796
Db        838 CGCCAGCGTGCCACCTGGTGGTCAGGGACTACGGGCACACGGACATGATGGACGACGACA 897

Qy        797 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 856
Db        898 CGCCGGGCGCCAGGGGGATGCTCACGCGCACCATCTGCAGGAGCGGCGGGGCCAGGGCGC 957

Qy        857 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 916
Db        958 CCATGCGCCGCTTCGTGGCCGGCGCCACCGTCGCGTTCCTCAAGAAATGGGTGGCTGGGG 1017

Qy        917 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGG 958
Db       1018 ACGCCGCGGCGATGGACAGCATCACGGCGCGGCCGGACCAGG 1059


RESULT 15
AF337544
LOCUS       AF337544                1115 bp    mRNA    linear   PLN 22-OCT-2002
DEFINITION  Brassica oleracea chlorophyllase 1 mRNA, complete cds.
ACCESSION   AF337544
VERSION     AF337544.1  GI:24210532
KEYWORDS    .
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 1115)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Molecular characterization and differential regulation of three
            chlorophyllase genes, members of the novel serine esterase gene
            family, in broccoli
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1115)
  AUTHORS   Chen,H.-H., Chou,Y.-F. and Shaw,J.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2001) Institute of Botany, Academia Sinica,
            Taipei, Taiwan 11529, Republic of China
FEATURES             Location/Qualifiers
     source          1. .1115
                     /organism="Brassica oleracea"
                     /mol_type="mRNA"
                     /cultivar="Green king"
                     /db_xref="taxon:3712"
```

```
                     /tissue_type="floret"
     5'UTR           1. .6
     CDS             7. .981
                     /EC_number="3.1.1.14"
                     /note="BoCLH1"
                     /codon_start=1
                     /product="chlorophyllase 1"
                     /protein_id="AAN51933.1"
                     /db_xref="GI:24210533"
                     /translation="MAGKEDSETFFSAATPLAFELGSLPTTVIPADPSATDLTAPPKP
                     VIITSPTVAGTYPVVLFFHGFYLRNYFYSDVINHVASHGYIVVAPQLCKILPPGGQVE
                     VDDAGKVINWTSKNLKAHLPSSVNANGNYTALVGHSRGGKTAFAVALGHAATLDPSIK
                     FSALVGIDPVAGISKCIRTDPEILTYKPESFDLDMPVAVIGTGLGPKSNMLMPPCAPA
                     EVNHEEFYIECKATKGHFVAADYGHMDMLDDNLPGFVGFMAGCMCKNGKRKKSEMRSF
                     VGGIVVAFLKYSIWGEMSEIRQILKDPSVSPARLDPSPELEEASGYLV"
     3'UTR           982. .1115
ORIGIN

  Query Match           16.2%;  Score 183.4;  DB 15;  Length 1115;
  Best Local Similarity  55.7%;  Pred. No. 1.9e-34;
  Matches  437;  Conservative    0;  Mismatches  336;  Indels   12;  Gaps    4;

Qy        139 CCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAG 198
              ||||| | | | ||   | | || |||||||   | | | | | ||   |
Db        103 CCGTCGGCAACCGATTTGACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAACCGTC 162

Qy        199 GAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTAT 258
              |  |||  ||| || || ||  |  || ||||||| | |  |||    |||| ||||||
Db        163 GCCGGAACTTACCCCGTCGTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTTCTAC 222

Qy        259 TCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATAT 318
              ||| |  ||||    || ||  |||||||||||| |||  |  | || || ||| | |
Db        223 TCTGATGTTATTAACCACGTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCTTTGC 282

Qy        319 AGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGG 378
              | ||      || |     |  ||  | |     | |   |  | ||  | |||
Db        283 AAGATTTTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAACTGG 342

Qy        379 TTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTT 438
                ||   |  || ||       || |||     |||||  | ||      | |
Db        343 ACTTCGAAAAACCTCAAAGCTCACCTCCCAAGTTCAGTAAACGCTAATGGCAACTACACC 402

Qy        439 GCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTT 498
              || |||   || ||||||||||||||| |||||||||||||||| || |||    |
Db        403 GCACTCGTGGGTCATAGCCGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCACGCC 462

Qy        499 G---GGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGA 555
              |       | |     | ||| | ||  |  |  | || ||||||||||| |  |||
Db        463 GCAACACTAGACCCATCCATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGCAGGA 522

Qy        556 ACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGAC 615
              |   | ||| | | |  |||| |||     | ||  | |||   ||  | ||||| |||
Db        523 ATCAGCAAATGCATAAGAACCGATCCCGAAATCTTAACGTACAAACCGGAATCATTCGAC 582

Qy        616 CTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCA 675
              || ||||   ||  |  |  |||||||| |||| || ||   | |     | |||
Db        583 CTGGACATGCCGGTTGCA---GTGATCGGTACGGGTCTCGG---ACCGAAGAGTAACATG 636

Qy        676 TTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAA 735
              |  | ||||| || |||||  | | ||||||  ||    ||||| | |   || ||| |
Db        637 CTGATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAG 696

Qy        736 GGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGAT 795
              | | |     | ||||||| ||   |||||| || ||| |||| ||| | || ||| ||
Db        697 GCTACGAAGGGACATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATAAT 756

Qy        796 ACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAA---GAGAGG 852
                  ||  ||  |||   |   |   |||| ||||||||||| ||| ||    | | |
Db        757 TTGCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACGGTAAACGCAAAAAG 816

Qy        853 AGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAA 912
              ||    ||||| || || |||||||| |||| ||  | ||| | |||  |  | |
Db        817 AGTGAGATGAGAAGCTTTGTTGGTGGAATTGTGGTTGCGTTTCTAAAGTATAGTATATGG 876

Qy        913 GGAGA 917
                || ||
Db        877 GGTGA 881


Search completed: March 20, 2006, 16:11:53
Job time : 4016.27 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:18 ; Search time 5143.9 Seconds
(without alignments)
10323.556 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		622.2	54.8	652	1	AU237094	AU237094 AU237094
2		539	47.5	641	9	AF005802	AF005802 AF005802
3		525	46.3	624	9	AF106729	AF106729 AF106729
c 4		507.8	44.7	883	10	CL497740	CL497740 SAIL_646_
c 5		399	35.2	399	3	BP789041	BP789041 BP789041
c 6		398	35.1	413	1	AU228115	AU228115 AU228115
c 7		375	33.0	427	3	BP786753	BP786753 BP786753
8		347.6	30.6	481	6	CA992230	CA992230 HC0810 GI
9		341.6	30.1	494	6	CA992233	CA992233 HC0819 GI
10		325	28.6	749	8	CX076457	CX076457 UCRCS08_5
11		312.4	27.5	504	11	CR397401	CR397401 Arabidops
c 12		308.6	27.2	317	1	AV556714	AV556714 AV556714
13		299.2	26.4	542	7	CN736070	CN736070 26RDBNT_U
c 14		298	26.3	827	8	CX076456	CX076456 UCRCS08_5
c 15		287	25.3	842	8	DR932000	DR932000 EST112353
16		283.4	25.0	761	6	CA815004	CA815004 CA12EI202
17		261.2	23.0	492	11	CR397400	CR397400 Arabidops
18		259	22.8	613	1	AJ808610	AJ808610 AJ808610
19		258.8	22.8	790	6	CF440994	CF440994 EST677339
20		253.6	22.3	569	7	CV525268	CV525268 Mdlv4010d
c 21		236.4	20.8	811	7	CO985817	CO985817 GM89015B1
22		224	19.7	611	7	CN490540	CN490540 Mdfw2017i
23		220.2	19.4	510	7	CV525641	CV525641 Mdlv4012n
c 24		213	18.8	776	6	CA816983	CA816983 CA12EI202
25		206.2	18.2	848	8	DR533022	DR533022 WS0274.B2
c 26		205.6	18.1	858	10	CZ950599	CZ950599 261087 To
27		205.4	18.1	634	3	BP955021	BP955021 BP955021
28		198.6	17.5	818	8	DR449908	DR449908 WS00951.B
29		195.8	17.3	561	1	AW039135	AW039135 EST281108
c 30		195	17.2	724	3	BQ113489	BQ113489 EST599065
31		193.6	17.1	499	3	BI699790	BI699790 sag48b06.
32		191	16.8	588	6	CD826708	CD826708 BN25.064M
33		189.4	16.7	577	3	BQ113488	BQ113488 EST599064
34		184	16.2	649	8	DR095974	DR095974 STRR1_24_
35		183.8	16.2	865	8	DR015096	DR015096 STRS1_1_B
36		182	16.0	758	9	BH511470	BH511470 BOGUP29TF
c 37		178.6	15.7	628	3	BP957354	BP957354 BP957354
38		175.2	15.4	750	8	DR539554	DR539554 WS0279.B2
39		172.2	15.2	497	6	CF608740	CF608740 GERMO01_0
40		169	14.9	541	3	BM885893	BM885893 sam09d1f.
41		164.8	14.5	668	7	CV187182	CV187182 Mdlv2_401
42		162.6	14.3	523	3	BM885730	BM885730 sam04e03.
43		154.8	13.6	435	2	BE923513	BE923513 EST427282
44		154.6	13.6	715	8	DR534907	DR534907 WS02744.C
45		149.6	13.2	699	7	CO487862	CO487862 GQ0252.B7

ALIGNMENTS

RESULT 1
AU237094
LOCUS AU237094 652 bp mRNA linear EST 01-APR-2002
DEFINITION AU237094 RAFL15 Arabidopsis thaliana cDNA clone RAFL15-48-F08 5', mRNA sequence.
ACCESSION AU237094
VERSION AU237094.1 GI:19876263
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 652)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998).cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further details.
FEATURES Location/Qualifiers
source 1. .652
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL15-48-F08"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_lib="RAFL15"
/note="Site_1: BamHI; Site_2: SalI"
ORIGIN

```
  Query Match           54.8%;  Score 622.2;  DB 1;  Length 652;
  Best Local Similarity  99.1%;  Pred. No. 4e-146;
  Matches  646;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy        117 ATCTCGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAG-C 175
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          1 ATCTTGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCC 60

Qy        176 TGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTT 235
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         61 TGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATGCGGTGGTGATGCTCCTCCATGGTT 120

Qy        236 ACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCA 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCA 180

Qy        296 TCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAAT 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAAT 240

Qy        356 CAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAG 415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAG 300

Qy        416 TAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGT 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGT 360

Qy        476 TTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCG 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCG 420

Qy        536 GTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTT 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTT 480

Qy        596 ACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTG 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTG 540

Qy        656 GTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCT-CCCGGAGTGAATCACCGA 714
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        541 GTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCCGGAGTGAATCACCGA 600

Qy        715 GAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATT 766
              ||||||||||||||||||| |||||||||||||||||||| |||||||||||
Db        601 GAGTTCTTTCGGGAATGTCCAGGTCCAGCATGGCATTTCGTTNCGAAGGATT 652


RESULT 2
AF005802
LOCUS       AF005802                641 bp    DNA     linear   GSS 06-NOV-2000
DEFINITION  AF005802 Arabidopsis thaliana 188-4 Arabidopsis thaliana genomic
            clone 1884xa similar to EST with GenBank Accession Number R30366,
            genomic survey sequence.
ACCESSION   AF005802
VERSION     AF005802.1  GI:3387726
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Mathur,J., Szabados,L., Schaefer,S., Grunenberg,B., Lossow,A.,
            Jonas-Straube,E., Schell,J., Koncz,C. and Koncz-Kalman,Z.
  TITLE     Gene identification with sequenced T-DNA tags generated by
            transformation of Arabidopsis cell suspension
  JOURNAL   Plant J. 13 (5), 707-716 (1998)
   PUBMED   9681013
COMMENT     Contact: Koncz C
            Abteilung Genetische Grundlagen der Pflanzenzuechtung
            Max-Planck Institut fuer Zuechtungsforschung
            Carl von Linne weg 10, Cologne, D-50829, Germany
            Email: koncz@mpiz-koeln.mpg.de
            transgenic cell line was obtained by transformation with the T-DNA
            of pPCV6NFHyg Agrobacterium binary vector; the left border junction
            of T-DNA insertion 1884xa was isolated in E. coli after XbaI
            digestion and self-circularization of plant DNA; clone 1884xa
            carries a plant DNA fragment of 5.0 kb that extends from a XbaI
            site to the left-border junction of pPCV6NFHyg T-DNA tag; sequences
            of the left T-DNA border are excluded from the submission
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-1"
                     /db_xref="taxon:3702"
                     /clone="1884xa"
                     /cell_line="188-4"
                     /clone_lib="Arabidopsis thaliana 188-4"
ORIGIN

  Query Match           47.5%;  Score 539;  DB 9;  Length 641;
  Best Local Similarity  99.5%;  Pred. No. 4.3e-125;
  Matches  550;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        310 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 149

Qy        370 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 209

Qy        430 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 269

Qy        490 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 329

Qy        550 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 389

Qy        610 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 449

Qy        670 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 509

Qy        730 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        510 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 569

Qy        790 GATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAG 849
              ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        570 GATGATAC-NNAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAG 628

Qy        850 AGGAGACCAATGA 862
              |||||||||||||
Db        629 AGGAGACCAATGA 641


RESULT 3
AF106729
LOCUS       AF106729                624 bp    DNA     linear   GSS 06-NOV-2000
DEFINITION  AF106729 Arabidopsis thaliana Col-0 Arabidopsis thaliana genomic
            clone 031-2XA.PC3, genomic survey sequence.
ACCESSION   AF106729
```

```
VERSION     AF106729.1  GI:4588400
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Koncz,C., Szabados,L., Gruneberg,B. and Schaefer,S.
  TITLE     Gene identification with sequenced T-DNA tags in Arabidopsis
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Koncz C
            Abteilung Genetische Grundlagen der Pflanzenzuechtung
            Max-Planck Institut fuer Zuechtungsforschung
            Carl von Linne weg 10, Cologne, D-50829, Germany
            Email: koncz@mpiz-koeln.mpg.de
            T-DNA tagged line 031-2, PCR fragment 031-2XA; Transgenic plant
            line 031-2 was obtained by transformation with the T-DNA of
            pPCV6NFHyg Agrobacterium binary vector; the left border junction of
            T-DNA insertion 031-2XA was isolated by PCR amplification using the
            LB2 (5'-CTGGGAATGGCGAAATCAAGGCATCGATCGTGAAG-3') and PC3
            (5'-CCTTGCGCCCTGAGTGCTTGCGGCAGC-3') primers and sequenced with the
            PC3 primer as described by Mathur et al. Plant J. (1998) 13,
            707-716; the PCR-amplified DNA fragment extends from an XbaI site
            to the left border junction of pPCV6NFHyg T- DNA tag, sequences of
            the left T-DNA end are excluded from the submission
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="031-2XA.PC3"
                     /clone_lib="Arabidopsis thaliana Col-0"
ORIGIN

  Query Match             46.3%;  Score 525;  DB 9;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-121;
  Matches  525;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        310 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 CAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATT 159

Qy        370 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 ATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTA 219

Qy        430 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 TCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTA 279

Qy        490 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 AAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTC 339

Qy        550 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCA 399

Qy        610 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 TTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGG 459

Qy        670 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 AACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAA 519

Qy        730 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 TGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGAT 579
```

```
Qy        790 GATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGT 834
              |||||||||||||||||||||||||||||||||||||||||||||
Db        580 GATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGT 624


RESULT 4
CL497740/c
LOCUS       CL497740                883 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  SAIL_646_E09.v1 SAIL Collection Arabidopsis thaliana genomic clone
            SAIL_646_E09.v1, genomic survey sequence.
ACCESSION   CL497740
VERSION     CL497740.1  GI:45992433
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 883)
  AUTHORS   Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
            Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
            Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
            Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
  TITLE     A high-throughput Arabidopsis reverse genetics system
  JOURNAL   Plant Cell 14 (12), 2985-2994 (2002)
  PUBMED    12468722
COMMENT     Contact: Sessions A
            Applied Trait Genetics
            Syngenta Biotechnology Inc.
            3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
            Email: allen.sessions@syngenta.com
            ABRC Stock Number CS827897; T-DNA left border flanking sequences of
            Syngenta Arabidopsis Insertion Library (SAIL) lines are available
            through the Arabidopsis Biological Resource Center (ABRC).
            Sequences represent a pool of amplified genomic regions and not
            single contiguous sequences.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .883
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SAIL_646_E09.v1"
                     /clone_lib="SAIL Collection"
                     /note="T-DNA left border sequences were isolated using a
                     modified TAIL-PCR strategy"
ORIGIN

  Query Match             44.7%;  Score 507.8;  DB 10;  Length 883;
  Best Local Similarity   96.8%;  Pred. No. 3.6e-117;
  Matches  518;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        351 TAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGC 410
              ||||||| || |||||||||||||||||||||||||| ||||||| ||||||||||||
Db        671 TAAATCATCGTCGGAGATTATGGATTGGTTATCAGTAGGCCTTAATCTCTTTCTTCCAGC 612

Qy        411 GCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAAC 470
              ||||||  | |||||||||||||||||||||||||||||| |||||||||||| |
Db        611 GCAAGTGGCCACAAACCTATCCAAATTTGCCCTCTCCGGCCATACCCGCGGTGGCAAAGC 552

Qy        471 CGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATT 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 CGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATT 492

Qy        531 GATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTT 590
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        491 GATCGGTATAGATCCGGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTT 432

Qy        591 GGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGG 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 GGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGG 372
```

```
Qy        651 GCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCA 710
              |||||||||| |||||||||||||||||||| |||||||| ||||||||||||||||||
Db        371 GCTTGGTGAACCCGCTCGGAACCCATTATTCCCGCCGTGTGCTCCTCCCGGAGTGAATCC 312

Qy        711 CCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGG 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 CCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGG 252

Qy        771 GCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTT 830
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        251 GCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATCGTTT 192

Qy        831 GTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTT 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        191 GTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTATT 137


RESULT 5
BP789041/c
LOCUS       BP789041                399 bp    mRNA    linear   EST 10-FEB-2005
DEFINITION  BP789041 RAFL7 Arabidopsis thaliana cDNA clone RAFL07-36-K05 3',
            mRNA sequence.
ACCESSION   BP789041
VERSION     BP789041.1  GI:59268282
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 399)
  AUTHORS   Seki,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K.,
            Iida,K., Enju,A., Sakurai,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Iida,J., Kawai,J., Sasaki,D., Shiraki,T., Hayashizaki,Y. and
            Shinozaki,K.
  TITLE     Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998, 2002).This clone is in a
            modified pBluescript vector.
            Please visit our web site (http://pfgweb.gsc.riken.jp and
            http://rarge.gsc.riken.jp) for further details.
            reversed clone.
FEATURES             Location/Qualifiers
     source          1. .399
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL07-36-K05"
                     /dev_stage="rosette plants"
                     /lab_host="DH10B"
                     /clone_lib="RAFL7"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     cold-treated (1, 2, 5, 10, 24 hr)"
ORIGIN

  Query Match            35.2%;  Score 399;  DB 3;  Length 399;
  Best Local Similarity  100.0%;  Pred. No. 1e-89;
  Matches  399;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        673 CCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGT 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 CCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGT 340

Qy        733 CAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGAT 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 CAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGAT 280

Qy        793 GATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGG 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 GATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGG 220

Qy        853 AGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAA 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 AGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAA 160

Qy        913 GGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAA 100

Qy        973 ATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 ATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTA 40

Qy       1033 TTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACA 1071
              ||||||||||||||||||||||||||||||||||||||
Db         39 TTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACA 1


RESULT 6
AU228115/c
LOCUS       AU228115                413 bp    mRNA    linear   EST 23-APR-2002
DEFINITION  AU228115 RAFL15 Arabidopsis thaliana cDNA clone RAFL15-48-F08 3',
            mRNA sequence.
ACCESSION   AU228115
VERSION     AU228115.1  GI:19742762
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
            Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
            Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998).cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL15-48-F08"
                     /tissue_type="mixture of silique and flower"
                     /lab_host="DH10B"
                     /clone_lib="RAFL15"
                     /note="Site_1: BamHI; Site_2: SalI"
ORIGIN

  Query Match            35.1%;  Score 398;  DB 1;  Length 413;
  Best Local Similarity  98.8%;  Pred. No. 1.8e-89;
  Matches  401;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy        715 GAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCAT 774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 GAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCAT 354

Qy        775 TTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGT 834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 TTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGT 294

Qy        835 AAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTT 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 AAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTT 234

Qy        895 TTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCAC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 TTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCAC 174

Qy        955 GAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTT 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTT 114

Qy       1015 TAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAAC 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 TAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAAC 54

Qy       1075 TTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAA 1120
              ||||||||||||||||||||||||||||||||||||  | |  ||
Db          53 TTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTATGATATTAA 8
```

RESULT 7
BP786753/c

```
LOCUS       BP786753                427 bp    mRNA    linear   EST 10-FEB-2005
DEFINITION  BP786753 RAFL7 Arabidopsis thaliana cDNA clone RAFL26-01-I24 3',
            mRNA sequence.
ACCESSION   BP786753
VERSION     BP786753.1  GI:59263657
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   Seki,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K.,
            Iida,K., Enju,A., Sakurai,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Iida,J., Kawai,J., Sasaki,D., Shiraki,T., Hayashizaki,Y. and
            Shinozaki,K.
  TITLE     Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998, 2002).This clone is in a
            modified pBluescript vector.
            Please visit our web site (http://pfgweb.gsc.riken.jp and
            http://rarge.gsc.riken.jp) for further details.
            reversed clone.
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL26-01-I24"
                     /dev_stage="rosette plants"
                     /lab_host="DH10B"
                     /clone_lib="RAFL7"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     cold-treated (1, 2, 5, 10, 24 hr)"
```

ORIGIN

```
 Query Match             33.0%;  Score 375;  DB 3;  Length 427;
  Best Local Similarity   98.7%;  Pred. No. 1.2e-83;
  Matches  378;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        738 TCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATAC 797
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 TCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATAC 368

Qy        798 AAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACC 857
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 AAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACC 308

Qy        858 AATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 AATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 248

Qy        918 TGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCA 977
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCA 188

Qy        978 AGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTC 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTC 128

Qy       1038 AATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGT 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGT 68

Qy       1098 CACCTCTTTGCTTACAAAAAAAA 1120
              ||||||||||||||  | |  ||
Db          67 CACCTCTTTGCTTATGATATTAA 45
```

RESULT 8
CA992230

```
LOCUS       CA992230                481 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  HC0810 GIBCOBRL CAT. NO. 19643-014 Brassica rapa subsp. pekinensis
            cDNA, mRNA sequence.
ACCESSION   CA992230
VERSION     CA992230.1  GI:37621525
KEYWORDS    EST.
SOURCE      Brassica rapa subsp. pekinensis
  ORGANISM  Brassica rapa subsp. pekinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 481)
  AUTHORS   Gao,R.J., Dai,D.P. and Ma,R.C.
  TITLE     Expressed sequence tags of heading leaf during the heading process
            of Chinese cabbage
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rongcai,Ma
            Plant Functional Genomics
            Beijing AgroBiotechnology Research Center
            Haidian District, Ban-Jing Rd., Beijing 100089, China
            Tel: 86 10 5150 3831
            Fax: 86 10 5150 3980
            Email: rcma1@yahoo.com
            Seq primer: T7
            High quality sequence stop: 481
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .481
                     /organism="Brassica rapa subsp. pekinensis"
                     /mol_type="mRNA"
                     /sub_species="pekinensis"
                     /db_xref="taxon:51351"
                     /tissue_type="heading leaf"
                     /dev_stage="folding stage"
                     /lab_host="E.coli DH10B(ZIP)"
                     /clone_lib="GIBCOBRL CAT. NO. 19643-014"
```

```
                     /note="Vector: pZL1; Site_1: NotI; Site_2: SalI; cDNA
                     library was constructed by SUPERSCRIPTTM Lambda System for
                     cDNA Synthesis and Cloning (GIBCOBRL,CAT.NO.19643-014)"
ORIGIN

  Query Match             30.6%;  Score 347.6;  DB 6;  Length 481;
  Best Local Similarity   84.3%;  Pred. No. 1e-76;
  Matches  403;  Conservative    0;  Mismatches   74;  Indels    1;  Gaps    1;

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567
              ||| | |||||||||||  || | |||||| ||||||| || ||||| ||||||||||
Db          2 TCGGAACTAAAGATCTCAGCACTTATCGGTGTAGATCCGGTGGATGGGACAGGGAAAGGT 61

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627
              |||||| |||||||||||||| ||   | ||    |||||||||||| ||||||| || | |
Db         62 AAACAAGCCCCTCCTCCGGTTTTAACCTATGAACCAAACTCATTTAACCTAGAAAAGATG 121

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687
              ||| | || || || |||||||| |||||||||   || |||||||||||| || ||||||
Db        122 CCTGTTCTAGTTATTGGTTCGGGACTTGGTGAACTTGCCCGGAACCCATTGTTTCCACCG 181

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747
              ||||||||| | |||||||| |||||||||||| || | |||||||||||||||||| |||
Db        182 TGTGCACCTACGGGAGTGAACCACCGAGAGTTTTTCCAGGAATGTCAAGGTCCAGCGTGG 241

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807
              ||||| || || |||||||| || ||||| ||||||||||||||||||||||||||||| ||
Db        242 CATTTTGTCGCAAAGGATTACGGACATTTAGACATGCTTGATGATGATACAAAAGGGCTT 301

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGA 867
              |||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||||
Db        302 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGAAAACCAATGAGGAGA 361

Qy        868 TTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAA 927
              ||  |||||||  |||||||  || |||||||  |||||||  |||||    ||||||  ||||
Db        362 TTTATTGGTGGTATTGTTGTGTCGTTTTTGATGGCTTATCTGGAAAATGATGATTGTGAG 421

Qy        928 TTAGTTAAGATCAAAGAT-GGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTT 984
              ||  | |||||||||  || ||||||||  || | ||||||  |||||||||||||  |||||
Db        422 TTGATGAAGATCAAAAATGGGGTGTCATGAAGGTGTTCCTGTTGAAATTCAAAAGTTT 479


RESULT 9
CA992233
LOCUS       CA992233                494 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  HC0819 GIBCOBRL CAT. NO. 19643-014 Brassica rapa subsp. pekinensis
            cDNA, mRNA sequence.
ACCESSION   CA992233
VERSION     CA992233.1  GI:37621528
KEYWORDS    EST.
SOURCE      Brassica rapa subsp. pekinensis
  ORGANISM  Brassica rapa subsp. pekinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Gao,R.J., Dai,D.P. and Ma,R.C.
  TITLE     Expressed sequence tags of heading leaf during the heading process
            of Chinese cabbage
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rongcai,Ma
            Plant Functional Genomics
            Beijing AgroBiotechnology Research Center
            Haidian District, Ban-Jing Rd., Beijing 100089, China
            Tel: 86 10 5150 3831
            Fax: 86 10 5150 3980
            Email: rcma1@yahoo.com
            Seq primer: T7
            High quality sequence stop: 494
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Brassica rapa subsp. pekinensis"
                     /mol_type="mRNA"
                     /sub_species="pekinensis"
                     /db_xref="taxon:51351"
                     /tissue_type="heading leaf"
                     /dev_stage="folding stage"
                     /lab_host="E.coli DH10B(ZIP)"
                     /clone_lib="GIBCOBRL CAT. NO. 19643-014"
                     /note="Vector: pZL1; Site_1: NotI; Site_2: SalI; cDNA
                     library was constructed by SUPERSCRIPTTM Lambda System for
                     cDNA Synthesis and Cloning (GIBCOBRL,CAT.NO.19643-014)"
ORIGIN

  Query Match             30.1%;  Score 341.6;  DB 6;  Length 494;
  Best Local Similarity   84.1%;  Pred. No. 3.4e-75;
  Matches  397;  Conservative    0;  Mismatches   74;  Indels    1;  Gaps    1;

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGG-TATAGATCCAGTCGATGGAACAGGGAAAGG 566
              ||| | |||||||||||  || | ||||| | ||||||| || ||||| ||||||||||
Db          2 TCGGAACTAAAGATCTCAGCACTTATCGGTTGTAGATCCGGTGGATGGGACAGGGAAAGG 61

Qy        567 GAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAAC 626
              |||||| |||||||||||||| ||  | ||    |||||||||||| ||||||| || |
Db         62 TAAACAAGCCCCTCCTCCGGTTTTAACCTATGAACCAAACTCATTTAACCTAGAAAAGAT 121

Qy        627 GCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACC 686
              |||| | || || || |||||||| |||||||||   || |||||||||||| || |||||
Db        122 GCCTGTTCTAGTTATTGGTTCGGGACTTGGTGAACTTGCCCGGAACCCATTGTTTCCACC 181

Qy        687 GTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATG 746
              |||||||||| | |||||||| |||||||||||| || | |||||||||||||||||| ||
Db        182 GTGTGCACCTACGGGAGTGAACCACCGAGAGTTTTTCCAGGAATGTCAAGGTCCAGCGTG 241

Qy        747 GCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGAT 806
              |||||| || || |||||||| || ||||| |||||||||||||||||||||||||||| |
Db        242 GCATTTTGTCGCAAAGGATTACGGACATTTAGACATGCTTGATGATGATACAAAAGGGCT 301

Qy        807 TAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAG 866
              |||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db        302 TAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGAAAACCAATGAGGAG 361

Qy        867 ATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGA 926
              |||  |||||||  |||||||  || |||||||  |||||||  ||||||  ||||||  ||||
Db        362 ATTTATTGGTGGTATTGTTGTGTCGTTTTTGATGGCTTATCTGGAAGATGATGATTGTGA 421

Qy        927 ATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAA 978
              ||  | |||||||||  |||||||||||  ||   || ||| ||||||||||||
Db        422 GTTGATGAAGATCAAAAATGGGTGTCATGAAAGTGGTCCTGTTGAAATTCAA 473


RESULT 10
CX076457
LOCUS       CX076457                749 bp    mRNA    linear   EST 14-DEC-2004
DEFINITION  UCRCS08_50B04_g Parent Washington Navel Orange Callus cDNA Library
            UCRCS08-2 Citrus sinensis cDNA clone UCRCS08-50B04-C8-1-5.g, mRNA
            sequence.
ACCESSION   CX076457
VERSION     CX076457.1  GI:56590447
KEYWORDS    EST.
SOURCE      Citrus sinensis
  ORGANISM  Citrus sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Close,T.J., Roose,M.L., Ye,X.R., Fenton,R.D., Wanamaker,S.,
            Lyon,M., Jang,C., Quinitio,C., Ikeda,J., Collin,M., Kacar,Y.,
            Landry,B., Hubert,N., Laforest,M., Landry,J. and Ligonde,A.
  TITLE     Development of EST Resources and New Genetic Markers for California
            Citrus - Washington Navel Orange Callus - UCRCS08-2
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Timothy J. Close
            Department of Botany & Plant Sciences
            University of California
```

Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T3.

```
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Citrus sinensis"
                     /mol_type="mRNA"
                     /cultivar="Washington navel"
                     /db_xref="taxon:2711"
                     /clone="UCRCS08-50B04-C8-1-5.g"
                     /tissue_type="Callus"
                     /dev_stage="Embryogenic and embryoid"
                     /lab_host="E. coli TJC121"
                     /clone_lib="Parent Washington Navel Orange Callus cDNA
                     Library UCRCS08-2"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Parent Washington navel
                     orange embryogenic callus was established from undeveloped
                     ovules of ca. 10 mm diameter young fruits under open
                     pollination on Murashige-Skoog medium at 25oC with 16 h
                     light in a tissue culture room. Embryogenic callus,
                     globular and heart stage embryoids were pooled in
                     approximately equal portions in RNAlater (Ambion), then
                     RNA was extracted using TRIZOL Reagent (Invitrogen).
                     Poly(A) RNA was purified from 500 microgram of total RNA
                     using Qiagen Oligotex. A primary cDNA library was produced
                     using a lambda ZAP XR cDNA Synthesis Kit (Stratagene).
                     These steps were performed by Xinrong Ye (Roose lab, UC
                     Riverside). One million pfu from the primary library were
                     mass excised to produce a phagemid population by Raymond
                     Fenton (Close lab, UC Riverside). Phagemids were plated,
                     plasmid DNA purified, cDNA clones archived, and DNA
                     sequences determined bi-directionally using an ABI3730 at
                     DNA Landmarks (Landry, Hubert, Laforest, Landry, Ligonde).
                     Chromatogram files were downloaded by FTP to UC Riverside
                     (by Close), then processed at UC Riverside (by Wanamaker,
                     Close lab) using the HarvEST pipeline
                     (http://harvest.ucr.edu) to remove vector and cloning
                     oligo sequences and various contaminants, and to trim to a
                     high quality region. Sequences that retained a phred 17
                     region of at least 100 bases were assembled, then chimeras
                     were removed following manual inspection of assemblies
                     (Close, Roose, Federici, Wanamaker, Lyon, Ye, Jang,
                     Collin, Kacar, Ikeda, Quinitio). Sequences that survived
                     all removal steps were submitted to GenBank."
ORIGIN

  Query Match             28.6%;  Score 325;  DB 8;  Length 749;
  Best Local Similarity   66.7%;  Pred. No. 5.8e-71;
  Matches  498;  Conservative    0;  Mismatches  240;  Indels    9;  Gaps    2;

Qy        127 TGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCT 186
              ||||| |  |||   ||  ||  |  | | ||  |   ||||      | |     || ||
Db         12 TGCAACAGTACAAGCTCCACTCCACTTCCGCCACCCAAGCCACTTTTGATCGGCATGCCT 71

Qy        187 ACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTAC 246
              |    |  |  |||||| |  || ||  ||||  | ||||| ||||||  |||||  |
Db         72 TCTGATGATGCTGGAGGAGAATTCCCAGTACTGATTTTACTCCACGGTTACGTTCTCCTC 131

Qy        247 AACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCT 306
              ||||| ||||| |||||||||| || |||||  |||||||||||||||  | ||||||
Db        132 AACTCTTTCTACTCTCAGCTTATCTTACATGTTGCTTCTCATGGCTTCATTGTAATCGCT 191

Qy        307 CCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAG 366
              |||||||||||  | | | ||  |||||||  |||  | ||||||||  ||  | || |
Db        192 CCTCAGTTATACAATGTGGCTGGACCAGATGCAACTGCTGAGATTACGTCTGCAGCTGCA 251

Qy        367 ATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAAC 426
              || |  |||||||||| | ||||||    ||||| |||||| | || ||    |||||
Db        252 ATCACAAATTGGTTATCTGAAGGACTCGGCCACTTCCTTCCACCACATGTTCGGCCAAAT 311

Qy        427 CTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCC 486
              |||  |||  | ||||||  | || |||||  || || ||||||  | || ||||||  | ||
Db        312 CTAAGCAAGCTAGCCCTTGCTGGTCATAGTCGAGGAGGCAAGGCTGCCTTTGCACTAGCT 371

Qy        487 TTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCA 546
              | ||||||  |      || | | | ||||    ||  |||| || ||| |||| ||
Db        372 CTTAAGAAAGGGG------CCACTACTTTAAAATATTCAGCATTAATTGGTGTAGACCCT 425

Qy        547 GTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAAC 606
              || |||||||  |  |||||||||||||| ||||| |||||  |  ||||| |||| |
Db        426 GTTGATGGAATGGACAAAGGGAAACAAACTCCTCCACCGGTACTAACTTACATTCCTCAT 485

Qy        607 TCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCT 666
              |||||||| || |    | |||  || | || || ||||| ||  | |||||||
Db        486 TCATTTGATCTTGGT---ATGCCGGTAATGGTTATAGGTTCAGGCTTAGGTGAAATTAAA 542

Qy        667 CGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGG 726
              |||||  | ||||| || ||||| ||    ||||| || |||  ||| |||||
Db        543 AAAAACCCTCTGTTCCCTCCTTGTGCTCCGAAAGGAGTTAACCACAAAGACTTCTTCAAC 602

Qy        727 GAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTT 786
              |||||||     || || || ||||| || |  ||||||||||| ||  | || ||| |
Db        603 GAATGTCGGACACCGGCCTGTCATTTTGTGGTTAAGGATTATGGTCACCTTGATATGTTG 662

Qy        787 GATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAA 846
              || |||||||| || ||||| || |||||      | || ||| ||||||||||||||| |
Db        663 GACGATGATACTAAGGGGATCAGGGGGAAAGCAACATACTGTCTGTGTAAGAATGGTAAG 722

Qy        847 GAGAGGAGACCAATGAGGAGATTCGTT 873
              |||  ||| |||||||| | | ||
Db        723 TCCAGGGAACCTATGAGGAGGTCCATT 749
```

RESULT 11
CR397401

```
LOCUS       CR397401                 504 bp    DNA     linear   GSS 02-MAY-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-461A10-026076,
            genomic survey sequence.
ACCESSION   CR397401
VERSION     CR397401.1  GI:46938129
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 504)
  AUTHORS   Strizhov,N., Rosso,M.G., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
```

```
          It indicates an insertion close to or within gene At5g43860.
          Details on the protocols used for generation of the sequence are
          described in References 1-3. The sequences are generated at the MPI
          for Plant Breeding Research in the context of the GABI-Kat project.
          GABI-Kat is part of the German Plant Genomics program designated
          'GABI'. Information on line availability can be found at:
          http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .504
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /clone="GK-461A10-026076"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match            27.5%;  Score 312.4;  DB 11;  Length 504;
  Best Local Similarity  99.7%;  Pred. No. 8.3e-68;
  Matches  313;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 174

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 234

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 294

Qy        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 354

Qy        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 414

Qy        301 ATCGCTCCTCAGTT 314
              |||||||||||| |
Db        415 ATCGCTCCTCAGGT 428


RESULT 12
AV556714/c
LOCUS       AV556714                317 bp    mRNA    linear   EST 23-FEB-2004
DEFINITION  AV556714 Arabidopsis thaliana green siliques Columbia Arabidopsis
            thaliana cDNA clone SQ050d08F 3', mRNA sequence.
ACCESSION   AV556714
VERSION     AV556714.1  GI:8728129
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 317)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7 (3), 175-180 (2000)
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .317
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /ecotype="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SQ050d08F"
                     /tissue_type="green siliques"
                     /clone_lib="Arabidopsis thaliana green siliques Columbia"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN

  Query Match            27.2%;  Score 308.6;  DB 1;  Length 317;
  Best Local Similarity  98.7%;  Pred. No. 6.9e-67;
  Matches  311;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        797 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        317 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 258

Qy        857 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 198

Qy        917 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTC 976
              ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        197 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGAATGTTCCCGTTGAAATTC 138

Qy        977 AAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGT 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 AAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGT 78

Qy       1037 CAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAG 1096
              |||||||||||||||||||||||||||||||| |||||||||||| | ||||||||||||
Db         77 CAATATCATCAGCTTTTGTTGCTTATGGTTTTGCAAACTTATATTGAAAAACTCTTTAAG 18

Qy       1097 TCACCTCTTTGCTTA 1111
              |||||||||||||||
Db         17 TCACCTCTTTGCTTA 3


RESULT 13
CN736070
LOCUS       CN736070                542 bp    mRNA    linear   EST 14-MAY-2005
DEFINITION  26RDBNT_UP_011_F07_20JAN2004_053 Brassica napus 26RDBNT Brassica
            napus cDNA 5', mRNA sequence.
ACCESSION   CN736070
VERSION     CN736070.1  GI:65293887
KEYWORDS    EST.
SOURCE      Brassica napus (rape)
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Xiang,D., Venglat,P., Keller,G., Selvaraj,G. and Datla,R.
  TITLE     Gene Expression Patterns during Brassica Zygotic Embryogenesis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Raju Datla
            Molecular and Developmental Genetics
            National Research Council of Canada
            Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon,
            Saskatchewan, S7N 0W9, Canada
            Tel: 306 975 5267
            Fax: 306 975 4839
            Email: Raju.Datla@nrc-cnrc.gc.ca
            High quality sequence stop: 542.
```

```
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Brassica napus"
                     /mol_type="mRNA"
                     /db_xref="taxon:3708"
                     /clone_lib="Brassica napus 26RDBNT"
                     /note="Vector: pDNR-LIB_CREATOR; Site_1: Sfi IA; Site_2:
                     Sfi IB; Zygotic embryos at torpedo stage from Brassica
                     napus seeds were used for the isolation of PolyA RNA and
                     in the construction of the cDNA library. Sequences have
                     been trimmed to remove vector and low quality regions
                     using LUCY sequence cleanup software (www.tigr.org)."
ORIGIN

  Query Match             26.4%;  Score 299.2;  DB 7;  Length 542;
  Best Local Similarity   74.8%;  Pred. No. 1.8e-64;
  Matches  389;  Conservative    0;  Mismatches  128;  Indels    3;  Gaps    1;

Qy        616 CTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCA 675
              |||||||||| |||| | || |||||||||||||| ||||||||||     | |||||
Db          1 CTAGACAAAATGCCTGTTCTAGTGATCGGTTCGGGACTTGGTGAAATTAGCCCGAACCTG 60

Qy        676 TTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAA 735
              ||||| || ||||| || |||||||||||||| |||||||| ||||| ||||||||||||
Db         61 TTATTTCCTCCGTGCGCGCCTCCCGGAGTGAACCACCGAGATTTCTTCCGGGAATGTCAA 120

Qy        736 GGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGAT 795
              |||||||| |||||||| || || |||||||| |||||||| ||||||||||||||||||
Db        121 GGTCCAGCGTGGCATTTTGTCGCAAAGGATTACGGGCATTTAGACATGCTTGATGATGAT 180

Qy        796 ACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGA 855
              ||| |||||||||||||| |||| |||| | | |||||||||||||||||||||||.|||
Db        181 ACAGAAGGGATTAGAGGCAAGATTTCTCACTATTTGTGTAAGAATGGTGAAGAGAGAAGA 240

Qy        856 CCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGA 915
              |||||||||||||| |||||||||  | ||||  | |||||||| ||||||||||||||||
Db        241 CCAATGAGGAGATTTGTTGGTGGAAGTATTGTGGCGTTTTTGATGGCTTATTTGGAAGGA 300

Qy        916 GATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATT 975
              |||||| ||||||| ||||| |||||||||  || |||||||| |||||| || ||||||
Db        301 GATGATAGTGAATTGGTTAAAATCAAAGATAAGTCTCACGAGGGTGTTCCTGTCGAAATT 360

Qy        976 CAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTG 1035
              || |||||||||||||  |  ||    |  |  | |||   |  ||     | | ||||
Db        361 CAGGAGTTTGAGGTTAAAAACTAGTTTTCTTTAGTGCTT---GTACTTTCCTATATATTC 417

Qy       1036 TCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAA 1095
              |  |||    |   | ||| |  | ||| | ||          | |  ||     ||||
Db        418 TACATAGTCGCTTTTGTTGCTTGTGATGTTATTTTGTGTGCTCTTGGCTCATGAAATTAA 477

Qy       1096 GTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                |      | |  |  |  ||| | | ||||||||||
Db        478 CCCCATCACATTCAAAAGACTAAAGAGACAAAAAAAAAAA 517


RESULT 14
CX076456/c
LOCUS       CX076456               827 bp    mRNA    linear   EST 14-DEC-2004
DEFINITION  UCRCS08_50B04_b Parent Washington Navel Orange Callus cDNA Library
            UCRCS08-2 Citrus sinensis cDNA clone UCRCS08-50B04-C8-1-4.b, mRNA
            sequence.
ACCESSION   CX076456
VERSION     CX076456.1  GI:56590446
KEYWORDS    EST.
SOURCE      Citrus sinensis
  ORGANISM  Citrus sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Close,T.J., Roose,M.L., Ye,X.R., Fenton,R.D., Wanamaker,S.,
            Lyon,M., Jang,C., Quinitio,C., Ikeda,J., Collin,M., Kacar,Y.,
            Landry,B., Hubert,N., Laforest,M., Landry,J. and Ligonde,A.
  TITLE     Development of EST Resources and New Genetic Markers for California
            Citrus - Washington Navel Orange Callus - UCRCS08-2
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Timothy J. Close
            Department of Botany & Plant Sciences
            University of California
            Riverside, CA 92521-0124, USA
            Tel: 909-787-3318
            Fax: 909-787-4437
            Email: timothy.close@ucr.edu
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Citrus sinensis"
                     /mol_type="mRNA"
                     /cultivar="Washington navel"
                     /db_xref="taxon:2711"
                     /clone="UCRCS08-50B04-C8-1-4.b"
                     /tissue_type="Callus"
                     /dev_stage="Embryogenic and embryoid"
                     /lab_host="E. coli TJC121"
                     /clone_lib="Parent Washington Navel Orange Callus cDNA
                     Library UCRCS08-2"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Parent Washington navel
                     orange embryogenic callus was established from undeveloped
                     ovules of ca. 10 mm diameter young fruits under open
                     pollination on Murashige-Skoog medium at 25oC with 16 h
                     light in a tissue culture room. Embryogenic callus,
                     globular and heart stage embryoids were pooled in
                     approximately equal portions in RNAlater (Ambion), then
                     RNA was extracted using TRIZOL Reagent (Invitrogen).
                     Poly(A) RNA was purified from 500 microgram of total RNA
                     using Qiagen Oligotex. A primary cDNA library was produced
                     using a lambda ZAP XR cDNA Synthesis Kit (Stratagene).
                     These steps were performed by Xinrong Ye (Roose lab, UC
                     Riverside). One million pfu from the primary library were
                     mass excised to produce a phagemid population by Raymond
                     Fenton (Close lab, UC Riverside). Phagemids were plated,
                     plasmid DNA purified, cDNA clones archived, and DNA
                     sequences determined bi-directionally using an ABI3730 at
                     DNA Landmarks (Landry, Hubert, Laforest, Landry, Ligonde).
                     Chromatogram files were downloaded by FTP to UC Riverside
                     (by Close), then processed at UC Riverside (by Wanamaker,
                     Close lab) using the HarvEST pipeline
                     (http://harvest.ucr.edu) to remove vector and cloning
                     oligo sequences and various contaminants, and to trim to a
                     high quality region. Sequences that retained a phred 17
                     region of at least 100 bases were assembled, then chimeras
                     were removed following manual inspection of assemblies
                     (Close, Roose, Federici, Wanamaker, Lyon, Ye, Jang,
                     Collin, Kacar, Ikeda, Quinitio). Sequences that survived
                     all removal steps were submitted to GenBank."
ORIGIN

  Query Match             26.3%;  Score 298;  DB 8;  Length 827;
  Best Local Similarity   67.4%;  Pred. No. 4e-64;
  Matches  453;  Conservative    0;  Mismatches  210;  Indels    9;  Gaps    2;

Qy        277 GTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGAC 336
              ||  |||||||||||||||  | |||||||||||||||| | | | || ||||||||
Db        827 GTTGCTTCTCATGGCTTCATTGTAATCGCTCCTCAGTTATACAATGTGGCTGGACCAGAT 768

Qy        337 ACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAAT 396
                |||  | ||||||||  ||  | || |  || |   ||||||||| | ||||||
Db        767 GCAACTGCTGAGATTACGTCTGCAGCTGCAATCACAAATTGGTTATCTGAAGGACTCGGC 708

Qy        397 CACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGC 456
              ||||| |||||| | || ||    ||||| |||  |||  | |||||  | || |||||
Db        707 CACTTCCTTCCACCACATGTTCGGCCAAATCTAAGCAAGCTAGCCCTTGCTGGTCATAGT 648

Qy        457 CGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTA 516
              || || |||||  | || |||||  | ||  | ||||||   |       || | | | ||
```

```
Db        647 CGAGGAGGCAAGGCTGCCTTTGCACTAGCTCTTAAGAAAGGGG------CCACTACTTTA 594

Qy        517 AAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACC 576

Db        593 AAATATTCAGCATTAATTGGTGTAGACCCTGTTGATGGAATGGACAAAGGGAAACAAACT 534

Qy        577 CCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTT 636

Db        533 CCTCCACCGGTACTAACTTACATTCCTCATTCATTTGATCTTGGT---ATGCCGGTAATG 477

Qy        637 GTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCT 696

Db        476 GTTATAGGTTCAGGCTTAGGTGAAATTAAAAAAAACCCTCTGTTCCCTCCTTGTGCTCCG 417

Qy        697 CCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTT 756

Db        416 AAAGGAGTTAACCACAAAGACTTCTTCAACGAATGTCGGACACCGGCCTGTCATTTTGTG 357

Qy        757 GCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAG 816

Db        356 GTTAAGGATTATGGTCACCTTGATATGTTGGACGATGATACTAAGGGGATCAGGGGGAAA 297

Qy        817 AGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGT 876

Db        296 GCAACATACTGTCTGTGTAAGAATGGTAAGTCCAGGGAACCTATGAGGAGGTCCATTGGA 237

Qy        877 GGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAG 936

Db        236 GGAATTATTGTTGCTTTTATGAAAGCTTATCTGGATGGTGATATCACCGACTTAATGGCT 177

Qy        937 ATCAAAGATGGG 948

Db        176 ATAAGAAAAGAG 165
```

RESULT 15
DR932000/c
LOCUS DR932000 842 bp mRNA linear EST 02-AUG-2005
DEFINITION EST1123539 Aquilegia cDNA library Aquilegia formosa x Aquilegia pubescens cDNA clone CO1OD44, mRNA sequence.
ACCESSION DR932000
VERSION DR932000.1 GI:71701363
KEYWORDS EST.
SOURCE Aquilegia formosa x Aquilegia pubescens
ORGANISM Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Aquilegia.
REFERENCE 1 (bases 1 to 842)
AUTHORS Hodges,S.A., Rensink,W., Buell,C.R., Borevitz,J., Kramer,E., Nordborg,M. and Tomkins,J.
TITLE Generation of ESTs from Aquilegia
JOURNAL Unpublished (2005)
COMMENT Other_ESTs: EST1123540
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: hodges@lifesci.ucsb.edu
Seq primer: TTTTTTTTTTTTTTTTTTTTTTTN (where N = A, G & C).
FEATURES Location/Qualifiers
source 1. .842
/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="CO1OD44"
/tissue_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site_1: EcoRI; Site_2: NotI; F2, F3, and F4 lines of Aquilegia formosa X A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

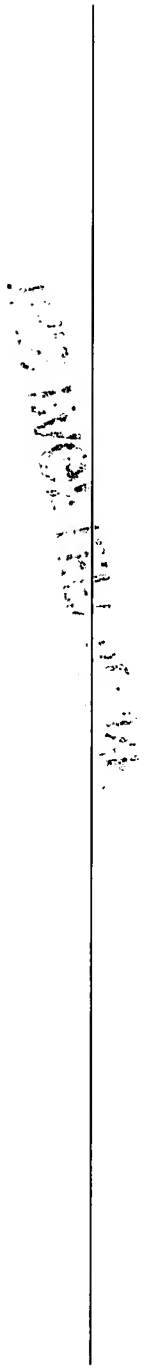
ORIGIN

Query Match 25.3%; Score 287; DB 8; Length 842;
Best Local Similarity 63.4%; Pred. No. 2.4e-61;
Matches 475; Conservative 0; Mismatches 265; Indels 9; Gaps 2;

```
Qy        196 GAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTC 255

Db        834 GAAGAGGGACAATTTCCAGTACTCATTTTTCTTCATGGCTTTCTTCTCTTTAATAAGTTT 775

Qy        256 TATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTA 315

Db        774 TATTCGCAATTAATTCAACACATCGCTTCCCATGGATTCATCGTTATTGCTCCTCAGTTA 715

Qy        316 TATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGAT 375

Db        714 TACAAAGTGGCAGGACCAGATACAACTGATGAGATAAAGTCAGCAGCTTTAGTAATAGAT 655

Qy        376 TGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAA 435

Db        654 TGGTTATCAAATGGCTTGCATAGTGTGCTTCCTCCACTAGTTCAACCAAATCTAAGCAAA 595

Qy        436 TTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAA 495

Db        594 CTAGGAATTGGAGGACATAGCAGGGGAGGCAAGGTTGCATTTGCACTGGCTCTTGGACAT 535

Qy        496 TTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGA 555

Db        534 ATTAAAACATCAT------TGAAATATTCGGTCTTGTTAGGTATAGATCCAGTTGATGGT 481

Qy        556 ACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGAC 615

Db        480 ATGGGCCAAGGAAATCAAACCCCTCCACCAGTTCTCACTTATACCCCTCGTTCTTTCGAT 421

Qy        616 CTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCA 675

Db        420 TT---CAACATGCCAGTCTTAGTTATTGGCTCTGGTTTGGGTGAAACAAAGAAAAACTCA 364

Qy        676 TTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAA 735

Db        363 CTCTTCCCACCTTGTGCTCCCAAGGGTGTAAACCACGAGAATTTCTATAGTGAGTGTTGT 304

Qy        736 GGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGAT 795

Db        303 TCACCGGCATGTTATTTTGTTGTAAAGGATTATGGTCATATGGACATGCTCGACGATGAT 244

Qy        796 ACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGA 855

Db        243 ACTGGAGGTGTACGAGGAAAAGCTACATATTGCACATGTAGCAATGGGAAAGCAAGGGAG 184

Qy        856 CCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGA 915

Db        183 CCAATGAGAACTTTTGTTGGAGGAATTATGGTTGCATTTATGAAAGCATATATGGAAAAT 124

Qy        916 GATGATCGTGAATTAGTTAAGATCAAAGA 944
```

Db 123 GACTCTCGTGATTTAATGGCTATTAAAGA 95

Search completed: March 20, 2006, 16:53:01
Job time : 5146.9 secs



THIS PAGE LEFT BLANK

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:51:15 ; Search time 667.906 Seconds
(without alignments)
11325.590 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1135	100.0	1135	5	AAH41137	Aah41137 Arabidops
2	1135	100.0	1135	12	ADJ98181	Adj98181 Thale cre
3	367.4	32.4	1216	6	AAD38075	Aad38075 Grape chl
4	256.8	22.6	1444	6	AAD38080	Aad38080 Tulip chl
5	237.4	20.9	987	6	AAD38074	Aad38074 Grape chl
6	201.6	17.8	1302	6	AAD38076	Aad38076 Corn chlo
7	184	16.2	1125	13	ADX51602	Adx51602 Plant ful
8	163.4	14.4	975	6	ABZ12353	Abz12353 Arabidops
9	163.4	14.4	975	6	ADG87791	Adg87791 A. thalia
10	163.4	14.4	975	6	ADG87590	Adg87590 A. thalia
11	163.4	14.4	1188	5	AAH41136	Aah41136 Arabidops
12	163.4	14.4	1188	12	ADJ98180	Adj98180 Thale cre
13	151	13.3	1174	6	AAD38077	Aad38077 Soybean c
14	148.2	13.1	1242	6	AAD38082	Aad38082 Wheat chl
15	145.4	12.8	1104	6	AAD38078	Aad38078 Soybean c
16	142.8	12.6	1125	6	AAD38079	Aad38079 Soybean c
17	136	12.0	1157	3	AAC34054	Aac34054 Arabidops
18	116	10.2	1244	13	ADX50647	Adx50647 Plant ful
19	116	10.2	1244	13	ADX50646	Adx50646 Plant ful
20	93.6	8.2	1340	5	AAH41135	Aah41135 Chenopodi
21	88.4	7.8	1156	6	AAD38086	Aad38086 Wheat chl
22	85.8	7.6	558	11	ACL32937	Acl32937 Rice abio
23	77.6	6.8	390	11	ACL27601	Acl27601 Rice abio
24	71.4	6.3	916	6	AAD38081	Aad38081 Wheat chl
25	64.8	5.7	1300	5	AAH41145	Aah41145 Chenopodi
26	52.2	4.6	1274	6	AAD38087	Aad38087 Corn chlo
c 27	49.2	4.3	2000	11	ACL37108	Acl37108 Rice stre
28	49	4.3	516	13	ACN51676	Acn51676 Cotton an
c 29	48	4.2	487	13	ACN48060	Acn48060 Cotton pr
30	47.4	4.2	467	13	ACN51760	Acn51760 Cotton an
c 31	47.4	4.2	552	13	ACN45312	Acn45312 Cotton pr
32	46.8	4.1	579	5	ABV29714	Abv29714 Human pro
33	46.8	4.1	579	5	ABV23838	Abv23838 Human pro
34	46.4	4.1	10039	6	ABL34043	Abl34043 Human imm
35	46	4.1	2422	3	AAC76970	Aac76970 Human ORF
c 36	45.8	4.0	419	8	ABX46069	Abx46069 Bovine ES
c 37	45.8	4.0	629	10	ABT21705	Abt21705 Breast ca
c 38	45.4	4.0	391	5	ABV08149	Abv08149 Human pro
c 39	45.4	4.0	626	5	ABV60941	Abv60941 Human pro
40	45.2	4.0	421	9	ACH19791	Ach19791 Human adu
41	45.2	4.0	464	13	ACN51998	Acn51998 Cotton an
c 42	45	4.0	351	4	AAL14805	Aal14805 Human bre
c 43	45	4.0	448	11	ACN84831	Acn84831 Breast ca
c 44	44.8	3.9	559	5	ABV07759	Abv07759 Human pro
45	44.8	3.9	2000	8	ADA71514	Ada71514 Rice gene

ALIGNMENTS

RESULT 1
AAH41137
ID AAH41137 standard; DNA; 1135 BP.
XX
AC AAH41137;
XX
DT 21-AUG-2001 (first entry)
XX
DE Arabidopsis thaliana chlorophyllase coding sequence #2.
XX
KW Chlorophyllase; transgenic plant; ds.
XX
OS Arabidopsis thaliana.
XX
PN JP2001086990-A.
XX
PD 03-APR-2001.
XX
PF 20-SEP-1999; 99JP-00266181.
XX
PR 20-SEP-1999; 99JP-00266181.
XX
PA (KAGO) KAGOME KK.
XX
DR WPI; 2001-338421/36.
DR P-PSDB; AAB99103.
XX
PT DNA encoding chlorophyllase, useful for producing transgenic plants.
XX
PS Claim 5; Page 12-14; 21pp; Japanese.
XX
CC The present sequence is a chlorophyllase protein coding sequence. The
CC chlorophyllase protein coding sequence can be used for the transformation
CC of a plant
XX
SQ Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 U; 0 Other;

Query Match 100.0%; Score 1135; DB 5; Length 1135;
Best Local Similarity 100.0%; Pred. No. 9.4e-283;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60

```
Db          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCA 60

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180

Qy        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240

Qy        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300

Qy        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360

Qy        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420

Qy        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480

Qy        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540

Qy        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600

Qy        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660

Qy        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720

Qy        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780

Qy        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840

Qy        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900

Qy        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960

Qy        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
```

RESULT 2
ADJ98181
ID ADJ98181 standard; DNA; 1135 BP.
XX
AC ADJ98181;
XX
DT 06-MAY-2004 (first entry)
XX
DE Thale cress chlorophyllase 2 DNA.
XX
KW phytol kinase; tocopherol biosynthesis; plant; drought resistance;
KW thale cress; ds; chlorophyllase 2.
XX
OS Arabidopsis thaliana.
XX
PN WO2004013312-A2.
XX
PD 12-FEB-2004.
XX
PF 05-AUG-2003; 2003WO-US025276.
XX
PR 05-AUG-2002; 2002US-0400689P.
PR 05-AUG-2003; 2003US-00634548.
XX
PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
PI Norris SR, Lincoln K, Abad MS, Eilers R, Hartsuyker KK;
PI Hirschberg J, Karunanandaa B, Moshiri F, Stein JC, Valentin HE;
PI Venkatesh TV;
XX
DR WPI; 2004-157125/15.
XX
PT New phytol kinase polynucleotides, useful in mediating tocopherol
PT biosynthesis and in producing plants with increased drought resistance.
XX
PS Claim 23; SEQ ID NO 19; 189pp; English.
XX
CC The invention relates to a novel substantially purified nucleic acid
CC molecule encoding a phytol kinase. The nucleic acid molecules and
CC polypeptides of the invention may be useful in mediating tocopherol
CC biosynthesis and in producing plants with increased drought resistance.
CC The current sequence is that of the thale cress chlorophyllase 2 DNA of
CC the invention.
XX
SQ Sequence 1135 BP; 331 A; 233 C; 247 G; 324 T; 0 U; 0 Other;

Query Match 100.0%; Score 1135; DB 12; Length 1135;
Best Local Similarity 100.0%; Pred. No. 9.4e-283;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCA 60

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180

Qy        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTT 240

Qy        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTC 300
```

```
Qy        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACG 360

Qy        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACA 420

Qy        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCG 480

Qy        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATA 540

Qy        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTT 600

Qy        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAA 660

Qy        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTC 720

Qy        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGAC 780

Qy        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAAT 840

Qy        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900

Qy        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGAT 960

Qy        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGG 1020

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
```

```
RESULT 3
AAD38075
ID   AAD38075 standard; cDNA; 1216 BP.
XX
AC   AAD38075;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Grape chlorophyllase cDNA from vrl1c.pk008.o21:fis clone.
XX
KW   Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Vitis sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .938
FT                   /*tag=  a
FT                   /product= "Grape chlorophyllase protein from
FT                   vrl1c.pk008.o21:fis clone"
FT                   /EC_number= "3.1.1.14"
FT                   /note= "CDS does not include start codon"
FT                   /partial
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23777.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 48-49; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is grape
CC   chlorophyllase cDNA from vrl1c.pk008.o21:fis clone
XX
SQ   Sequence 1216 BP; 331 A; 275 C; 256 G; 354 T; 0 U; 0 Other;
```

Query Match 32.4%; Score 367.4; DB 6; Length 1216;
Best Local Similarity 63.9%; Pred. No. 2.2e-84;
Matches 608; Conservative 0; Mismatches 331; Indels 12; Gaps 3;

```
Qy         50 CTTCTTCATCAAGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGG 109
              ||||| |||      | | |   || | |||      | |     ||| | | | ||  ||
Db          1 CTTCTACATGTTCCACCACTGCTGCAAATGTTTTTGAGATTGGAAAGCACATCACAGTGC 60

Qy        110 ACTCATCATCTCGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAA 169
                |  | ||        |||      ||  |    ||       | |     | || |||
Db         61 TTCTAAGAGCTGAACCAGGCACTTGCACCACCAAGTCATCTCTTCCTGTTCCCCCTCCAC 120

Qy        170 AGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCC 229
               ||| || ||| | ||||||||     || | ||| || |   ||| || || | ||||| |
Db        121 TGCAACTCTTGATTGCTACGCCCTCTGAAGCAGGGGAGTTCCCGCTGCTGCTCCTCCTTC 180

Qy        230 ATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATGTCTCTTCTCATG 289
              ||||||| |||||||| ||||| ||||| || ||||| ||    ||  |  | |||||||
Db        181 ATGGTTATCTTCTCTATAACTCTTTCTACTCCCAGCTCATCCAACACATAGCCTCTCATG 240

Qy        290 GCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACACAATGGATGAGA 349
              | |||||  |  | |||||||||||||| | | | || |||||||  |||  || ||||
Db        241 GTTTCATTGTTCTTGCTCCTCAGTTATACACTGTGGCTGGACCAGATTCAAGCGAAGAGA 300

Qy        350 TTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATCACTTTCTTCCAG 409
              | || ||  | || |   | |  ||||||||||   ||||||  || |||| |||||
Db        301 TCAAGTCCGCAGCTGCTTTAACAAATTGGTTATCCAAAGGACTCCATGACTTACTTCCTC 360

Qy        410 CGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAA 469
              | || ||   |||||  ||  |||| | |  ||  |||||||||| || || |||||||
Db        361 CCCATGTTCGGCCAAATTTAAGCAAACTAGGACTTGCCGGCCATAGTCGTGGAGGCAAAA 420
```

```
Qy        470 CCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAAAGATCTCGACAT 529
              | || |||||  | ||  | |  |||          ||| |   ||| ||  | ||  | |
Db        421 CTGCTTTTGCTCTAGCACTGAGAAAA------GCATCCACTTCTCTGAAATTTTCAGCCT 474

Qy        530 TGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGT 589
              |||| || ||||| || || |||||||  |  ||||||||||||||||||| |||||
Db        475 TGATAGGCATAGACCCGGTGGATGGAATGGACAAAGGGAAACAAACCCCTCCACCGGTAC 534

Qy        590 TGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTGTGATCGGTTCGG 649
              |  | ||  ||||  | ||||||||| ||||||   | | |  |  | || || |||||||
Db        535 TCACCTATGTTCCTCATTCATTTGATCTAGAC---ATGGCAGTGATGGTAATTGGTTCGG 591

Qy        650 GGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTCCCGGAGTGAATC 709
              |  | ||||||       |||||||  | ||||| || ||||| ||    || || || |
Db        592 GTTTGGGTGAAGTGAAAAGGAACCCTCTGTTCCCTCCTTGTGCCCCCAAGGGCGTAAACC 651

Qy        710 ACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATG 769
              |    || ||||||   ||||| |  |  ||||| ||  ||||  |||| ||||| ||||
Db        652 ATGAGGACTTCTTTAAAGAATGCCGTGAACCAGCTTGTTATTTTCTTGCCAAGGACTATG 711

Qy        770 GGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTT 829
              | ||| | |||||||| || ||||| || || || ||||||||||   | |  ||||||
Db        712 GCCATCTTGACATGCTAGACGATGAGACTAATGGAATTAGAGGGAAAGCTACACATTGTT 771

Qy        830 TGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTAT 889
              ||||||| ||||  |    ||  ||| |||||||| || ||||| ||  ||||  |
Db        772 TGTGTAAAAATGGGAAGTCTAGAGAACCCATGAGGAGGTTTGTTGGAGGCATTGTGATTG 831

Qy        890 CATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGT 949
              ||||| |||| |||||||||||||| ||| |  |     || |    || | |||||||
Db        832 CATTTATGAAAGCTTATTTGGAAGGCGATAACAGCAGTCTAATCTCCATTAGAGATGGG- 890

Qy        950 GTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAA 1000
                || |    ||  || || ||  ||||| |  ||||| || || ||| ||
Db        891 --CATGCTACTGCACCAGTGGAGCTTCAAAATGTTGAGTTTCTCGTGTGAA 939
```

RESULT 4
AAD38080
ID AAD38080 standard; cDNA; 1444 BP.
XX
AC AAD38080;
XX
DT 07-AUG-2003 (revised)
DT 10-SEP-2002 (first entry)
XX
DE Tulip chlorophyllase cDNA from etp1c.pk005.d16:fis clone.
XX
KW Tulip; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
OS Tulipa gesneriana.
XX
FH Key Location/Qualifiers
FT CDS 97. .1107
FT /*tag= a
FT /product= "Tulip chlorophyllase protein from
FT etp1c.pk005.d16:fis clone"
FT /EC_number= "3.1.1.14"
XX
PN WO200229022-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031059.
XX
PR 05-OCT-2000; 2000US-0238161P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Thorpe CJ;
XX
DR WPI; 2002-444102/47.
DR P-PSDB; AAE23782.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 56-57; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is tulip
CC chlorophyllase cDNA from etp1c.pk005.d16:fis clone. (Updated on 07-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 1444 BP; 460 A; 295 C; 298 G; 391 T; 0 U; 0 Other;

Query Match 22.6%; Score 256.8; DB 6; Length 1444;
Best Local Similarity 61.1%; Pred. No. 8.9e-56;
Matches 474; Conservative 0; Mismatches 287; Indels 15; Gaps 3;

```
Qy        154 TCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCG 213
              ||  |  | ||  |  ||||  |    ||  | ||  ||  ||   | ||||||| || || ||
Db        199 TCCGCACCACCTCGGAAGCCACCACTGATCGCAACTCCATTTGAGGAAGGCGAGTACCCA 258

Qy        214 GTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTG 273
                |  |  | ||||| ||||| | | | ||  ||||| | ||||| ||| |||||||   |
Db        259 ACGCTTCTACTCCTTCATGGATTCATGCTTCACAACACTTTCTACTCTGAGCTTATCCAG 318

Qy        274 CATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCA 333
              ||  || | || |||||||| ||   |  | |  ||||||||||||  | | ||   |
Db        319 CACATCGCATCCCATGGCTTTATTGTTGTTGTACCTCAGTTATATCTTGTAGCTACATGT 378

Qy        334 GACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTT 393
              || |  | | |||  || || ||  | ||  | |  |  ||||||||     |  |||||
Db        379 GATAGTACGAATGGCATCAAGTCTGCTGCAAAAACAACAGATTGGTTGAAGGATGGACTG 438

Qy        394 AATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCAT 453
              |  |  |||| ||| |  |||| | |||| |||||   ||| | |  ||   ||| |||
Db        439 CAAGATGTTCTCCCAACAAAAGTCAGACCAGACCTAAAGAAACTCGGACTGAGCGGACAT 498

Qy        454 AGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAAT 513
              ||||| || ||||||   || |||||  | ||  ||     |       |     | | |
Db        499 AGCCGTGGCGGCAAAGATGCATTTGCTCTTGCACTAGGATATGCGAAG------ACTACA 552

Qy        514 CTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAA 573
              |||   |||| |  |||||| || ||||| || || |||  | | || ||  | |||
Db        553 TTAAGCTTCTCAGCGCTGATCGGGATCGATCCTGTTGACGGAGTACGAAAGGGACACCAA 612

Qy        574 ACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATA 633
              |||  |||||| ||  ||  |||  | ||  |||| || || ||   ||| | |||  |
Db        613 ACCAATCCTCCTGTACTGAATTATATCCCTCACTCTTTGGAACT---CAAGATGCCGTCA 669

Qy        634 CTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCA 693
              | || |||||  | ||  | ||||||       |||||| |||   |      |||||
Db        670 TTAGTTATCGGAACAGGTTTAGGTGAATTGAAGAGGAACCTATTTGCC------TGTGCC 723

Qy        694 CCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTC 753
              |||   ||||| ||||||| ||| || |      || |||       ||  | || |||||
Db        724 CCTAAGGGAGTAAATCACCAAGATTTTTACGATGAGTGTTCTTCCCCGTCTTGCCATTTT 783

Qy        754 GTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGG 813
              ||||| ||||| ||||| ||  |||||||||| |||||| ||||  | ||   ||||||
Db        784 GTTGCTAAGGACTATGGTCACATGGACATGCTAGATGATCATACTGACGGACCTAGAGGT 843

Qy        814 AAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTT 873
              |      || || ||  |||||  ||||| |||   || | ||| ||| |    || ||
Db        844 ATTGTGTCATACTGCATGTGTCGAAATGGAGAATCAAGAACACCCATGCGACTGTTTGTA 903
```

```
Qy        874 GGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATT 929
Db        904 GGTGGGATCGTGGTTGCTTTTCTGAAAGCCTGCTTACAAGGTGATAAACAGGATTT 959

RESULT 5
AAD38074
ID   AAD38074 standard; cDNA; 987 BP.
XX
AC   AAD38074;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Grape chlorophyllase cDNA from vdb1c.pk002.p19:fis clone.
XX
KW   Grape; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Vitis sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..858
FT                   /*tag=  a
FT                   /product= "Grape chlorophyllase protein from
FT                   vdb1c.pk002.p19:fis clone"
FT                   /EC_number= "3.1.1.14"
FT                   /note= "CDS does not include start codon"
FT                   /partial
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23776.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 47; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is grape
CC   chlorophyllase cDNA from vdb1c.pk002.p19:fis clone
XX
SQ   Sequence 987 BP; 260 A; 224 C; 218 G; 285 T; 0 U; 0 Other;

  Query Match             20.9%;  Score 237.4;  DB 6;  Length 987;
  Best Local Similarity   56.3%;  Pred. No. 7.9e-51;
  Matches  551;  Conservative    0;  Mismatches  411;  Indels   17;  Gaps    5;

Qy        158 CGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGG 217
Db         23 CTTCCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTC 82

Qy        218 TGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATG 277
Db         83 TCTTGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCA 142

Qy        278 TCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACA 337
Db        143 TTTCTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTG 202

Qy        338 CAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATC 397
Db        203 GAATTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAAT 262

Qy        398 ACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCC 457
Db        263 CTGTGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCA 322

Qy        458 GCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAA 517
Db        323 GAGGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGCTGATACATCC------CTCA 376

Qy        518 AGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCC 577
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        578 CTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTG 637
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAG---CAATCCCAGTTTGCG 493

Qy        638 TGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTC 697
Db        494 TAATCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAG 553

Qy        698 CCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTG 757
Db        554 ATGGAGTGAACCATGTAGAGTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGA 613

Qy        758 CGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGA 817
Db        614 CTACTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGA 673

Qy        818 GTTCTTATTGTTTGTGTAAGAATGGTGAAG---AGAGGAGACCAATGAGGAGATTCGTTG 874
Db        674 TTTCGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGG 733

Qy        875 GTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTA 934
Db        734 GTGGCCTTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATT---TCA 790

Qy        935 AGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCA 994
Db        791 AAGCCATTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCTGTTGAGTTCATAG 850

Qy        995 TGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTG 1054
Db        851 AAGCATGAAATCATGGATTCTTTGTAATGCTATAAGGAGTGTGA--CAGCAGATTTTTTT 908

Qy       1055 TTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAA 1114
Db        909 TTCGGTTCGTCTTTGGGAACAAAGATTCAAATAAACGAAACTGTGTTGTATTGCTAAAAA 968

Qy       1115 AAAAAAAAAAAAAAAAAAA 1133
Db        969 AAAAAAAAAAAAAAAAAAA 987

RESULT 6
AAD38076
ID   AAD38076 standard; cDNA; 1302 BP.
XX
AC   AAD38076;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Corn chlorophyllase cDNA from csh3c.pk001.a9:fis clone.
XX
KW   Corn; chlorophyllase; chlorophyll degradation; plant cell senescence;
```

```
KW   enzyme; gene; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             72. .1112
FT                   /*tag=  a
FT                   /product= "Corn chlorophyllase protein from
FT                   csh3c.pk001.a9:fis clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23778.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 50; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is corn
CC   chlorophyllase cDNA from csh3c.pk001.a9:fis clone
XX
SQ   Sequence 1302 BP; 248 A; 424 C; 407 G; 223 T; 0 U; 0 Other;

  Query Match             17.8%;  Score 201.6;  DB 6;  Length 1302;
  Best Local Similarity   54.6%;  Pred. No. 1.6e-41;
  Matches  449;  Conservative    0;  Mismatches  364;  Indels    9;  Gaps    2;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
Db        241 CGTCCAGCGCGGCCAGCCTGCCGCCGAAGCCGCTGCTCGTCGCCGCGCCGCGCGAGACTG 300

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
Db        301 GGGAGTACCCGGTGATCCTGTTCCTACACGGCTACCTCGCCGTCAACTCCTTCTACTCCC 360

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
Db        361 AGCTGTTCGAGCACGTCGCCTCCCATGGCTTTATCGTCGTCGGACCTCAGCTGTACACCA 420

Qy        323 TCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTAT 382
Db        421 TATCTGGGGCCGACACCACCGAGGAGATCAACTCAGCGGCGGCCGTCATCGACTGGCTAG 480

Qy        383 CAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCC 442
Db        481 CCACCGGGCTGCCGTCAACTCTGCCACTCGGCGTCCGCGCGAACCTAACCAAGGTGTCCA 540

Qy        443 TCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGT 502
Db        541 TCTCCGGCCACAGTCGCGGCGGGAAGGTGGCGTTCGCGCTGGCGTTGGGCCACGCCAAGG 600

Qy        503 ACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGA 562
Db        601 CCAAGCTCGCTGTCCCTCTCGCCGCCGTCGTCGCCGTGGACCCGGTGGACGGCATGGGCG 660

Qy        563 AAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACA 622
Db        661 TGGGCAAGCAGACACCGCCGCCGATCCTCACGGGCAGGCACGGCTCGCTGCACGTGG--- 717

Qy        623 AAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCC 682
Db        718 GTGCCCCCGCCATGGTCATCGGCACGGGGCTCGGCGAGCTGCCCCGCGGCTCACTGCTCC 777

Qy        683 CACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCT------TTCGGGAATGTCAAG 736
Db        778 CGCCGTGCGCGCCCCGGGGCGTCAGTCACGCGGCCTTCTACGACGAGCTGGACGGCGCGG 837

Qy        737 GTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATA 796
Db        838 CGCCAGCGTGCCACCTGGTGGTCAGGGACTACGGGCACACGGACATGATGGACGACGACA 897

Qy        797 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 856
Db        898 CGCCGGGCGCCAGGGGGATGCTCACGCGCACCATCTGCAGGAGCGGCGGGGCCAGGGCGC 957

Qy        857 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 916
Db        958 CCATGCGCCGCTTCGTGGCCGGCGCCACCGTCGCGTTCCTCAAGAAATGGGTGGCTGGGG 1017

Qy        917 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGG 958
Db       1018 ACGCCGCGGCGATGGACAGCATCACGGCGCGGCCGGACCAGG 1059


RESULT 7
ADX51602
ID   ADX51602 standard; cDNA; 1125 BP.
XX
AC   ADX51602;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polynucleotide seqid 26342.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content; gene; ss.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
```

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 26342; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 1125 BP; 293 A; 260 C; 274 G; 298 T; 0 U; 0 Other;

Query Match 16.2%; Score 184; DB 13; Length 1125;
Best Local Similarity 56.1%; Pred. No. 5.2e-37;
Matches 431; Conservative 0; Mismatches 325; Indels 12; Gaps 4;

Qy 139 CCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAG 198
Db 135 CCGTCGGCAACCGATTTGACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAACCGTC 194

Qy 199 GAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTAT 258
Db 195 GCCGGAACTTACCCCGTCGTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTTCTAC 254

Qy 259 TCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATAT 318
Db 255 TCTGATGTTATTAACCACGTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCTGTGC 314

Qy 319 AGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGG 378
Db 315 AAGATTTTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAACTGG 374

Qy 379 TTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTT 438
Db 375 ACTTCGAAAAACCTCAAAGCTCACCTCCCAAGCTCAGTAAACGCTAATGGCAACTACACC 434

Qy 439 GCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTT 498
Db 435 GCACTCGTGGGCCATAGCCGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCACGCC 494

Qy 499 G---GGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGA 555
Db 495 GCAACACTAGACCCATCCATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGCAGGA 554

Qy 556 ACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGAC 615
Db 555 ATCAGCAAATGCATAAGAACCGATCCCGAAATCTTAACGTACAAACCGGAATCATTCGAC 614

Qy 616 CTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCA 675
Db 615 CTGGACATGCCGGTTGCA---GTGATCGGTACGGGTCTCGG---ACCGAAGAGTAACATG 668

Qy 676 TTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAA 735
Db 669 CTGATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAG 728

Qy 736 GGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGAT 795

Db 729 GCTACCAAGGGGCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATAAT 788

Qy 796 ACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAA---GAGAGG 852
Db 789 TTGCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACAGTAAACGCAAAAAG 848

Qy 853 AGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
Db 849 AGTGAGATGAGAAGCTTTGTTGGTGGAATTGTGGTTGCGTTTCTAAAG 896

RESULT 8
ABZ12353
ID ABZ12353 standard; DNA; 975 BP.
XX
AC ABZ12353;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 158.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 158; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

Query Match 14.4%; Score 163.4; DB 6; Length 975;
Best Local Similarity 54.3%; Pred. No. 1e-31;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

Qy 151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210
Db 109 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 168

Qy 211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

```
Db        169 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 228

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330

Db        229 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 288

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390

Db        289 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 348

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450

Db        349 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 408

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507

Db        409 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 468

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Db        469 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 528

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Db        529 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 588

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Db        589 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 642

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Db        643 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 702

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807

Db        703 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 762

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864

Db        763 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 822

Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917

Db        823 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 875
```

```
RESULT 9
ADG87791
ID   ADG87791 standard; cDNA; 975 BP.
XX
AC   ADG87791;
XX
DT   22-APR-2004  (first entry)
XX
DE   A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #233.
XX
KW   Pathogen infection-related gene; plant; Peronospora parasitica;
KW   defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW   oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200222675-A2.
XX
PD   21-MAR-2002.
XX
PF   14-SEP-2001; 2001WO-US028506.
XX
PR   15-SEP-2000; 2000US-0232778P.
PR   22-JUN-2001; 2001US-0300183P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA   (UYNC-) UNIV NORTH CAROLINA.
PA   (GLAZ/) GLAZEBROOK J.
PA   (WANG/) WANG X.
PA   (DANG/) DANGL J L.
PA   (EULG/) EULGEM T.
PA   (ZHUT/) ZHU T.
XX
PI   Glazebrook J,  Wang X,  Dangl JL,  Eulgem T,  Zhu T;
XX
DR   WPI; 2002-292409/33.
XX
PT   Novel isolated polynucleotide, useful for conveying pathogen resistance
PT   to plants, and for identifying plants infected with a pathogen.
XX
PS   Claim 3; SEQ ID NO 233; 605pp; English.
XX
CC   The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC   ADG87557}) whose expression is altered in response to pathogen infection,
CC   and to homologues of these genes from other plants or fungi, especially
CC   from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC   cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC   expression of genes of the invention was upregulated or downregulated in
CC   Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC   indicating that they play a role in defence mechanisms. The genes of the
CC   invention are regulated by RPP7 or RRP8 which act via unconventional
CC   signalling cascades, or by the RPP4-dependent pathway. The invention also
CC   relates to polypeptides encoded by the pathogen infection-related genes;
CC   promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC   ; expression cassettes, host cells and pathogen-resistant transgenic
CC   plants and their progeny comprising a polynucleotide of the invention;
CC   and a method of identifying a plant cell infected with a pathogen. The
CC   polynucleotide sequences and methods of the invention are useful for
CC   identifying plants infected with a pathogen, and for conferring
CC   resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC   neamotodes and insects (e.g., aphids). The present sequence represents an
CC   Arabidopsis thaliana gene whose expression is altered in response to
CC   Peronospora parasitica infection. Note: The sequence data for this patent
CC   did not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

  Query Match            14.4%;  Score 163.4;  DB 6;  Length 975;
  Best Local Similarity  54.3%;  Pred. No. 1e-31;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;

Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Db        109 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 168

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Db        169 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 228

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330

Db        229 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 288

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390

Db        289 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 348

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450

Db        349 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 408

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507

Db        409 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 468

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567
```

```
Db        469 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 528

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Db        529 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 588

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Db        589 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 642

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Db        643 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 702

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807

Db        703 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 762

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864

Db        763 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 822

Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917

Db        823 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 875
```

RESULT 10
ADG87590
ID ADG87590 standard; cDNA; 975 BP.
XX
AC ADG87590;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #32.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP7; RRP8; pathogen resistance; transgenic plant;
KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200222675-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANGL J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
PI Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX
DR WPI; 2002-292409/33.
XX
PT Novel isolated polynucleotide, useful for conveying pathogen resistance
PT to plants, and for identifying plants infected with a pathogen.
XX
PS Claim 3; SEQ ID NO 32; 605pp; English.
XX
CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557}) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RRP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC neamotodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 975 BP; 250 A; 222 C; 256 G; 247 T; 0 U; 0 Other;

Query Match 14.4%; Score 163.4; DB 6; Length 975;
Best Local Similarity 54.3%; Pred. No. 1e-31;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;

```
Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Db        109 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 168

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Db        169 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 228

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330

Db        229 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 288

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390

Db        289 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 348

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450

Db        349 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 408

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507

Db        409 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 468

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Db        469 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 528

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Db        529 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 588

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Db        589 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 642

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Db        643 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 702

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807

Db        703 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 762

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864
```

```
Db        763 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 822

Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
              || || || |||||| |||| ||  | ||| | |||  |  ||||   || ||
Db        823 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 875


RESULT 11
AAH41136
ID   AAH41136 standard; DNA; 1188 BP.
XX
AC   AAH41136;
XX
DT   21-AUG-2001  (first entry)
XX
DE   Arabidopsis thaliana chlorophyllase coding sequence #1.
XX
KW   Chlorophyllase; transgenic plant; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   JP2001086990-A.
XX
PD   03-APR-2001.
XX
PF   20-SEP-1999;   99JP-00266181.
XX
PR   20-SEP-1999;   99JP-00266181.
XX
PA   (KAGO ) KAGOME KK.
XX
DR   WPI; 2001-338421/36.
DR   P-PSDB; AAB99102.
XX
PT   DNA encoding chlorophyllase, useful for producing transgenic plants.
XX
PS   Claim 3; Page 10-11; 21pp; Japanese.
XX
CC   The present sequence is a chlorophyllase protein coding sequence. The
CC   chlorophyllase protein coding sequence can be used for the transformation
CC   of a plant
XX
SQ   Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;

  Query Match             14.4%;  Score 163.4;  DB 5;  Length 1188;
  Best Local Similarity   54.3%;  Pred. No. 1.1e-31;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;

Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210
              | |||| |  | |||||||| | | ||    |  |    ||     |  |  |||  ||||
Db        157 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 216

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270
              || || ||  |  || |||||||  |    |||   ||||| |||||| ||| |  || |
Db        217 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 276

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330
                ||  || |||| ||||| | ||| ||  | || || |||||  |  |   |
Db        277 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 336

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390
              || |        |  ||  |  |      | |       | ||  | |||  ||| | |
Db        337 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 396

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450
              || ||       || ||| |    ||||   | ||   |     |     | |||   |||
Db        397 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 456

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507
              || |||||||||||| ||||| |||||||||||| ||  ||    |    |     |   |
Db        457 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 516

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567
              |     | | | | ||  |  | || || || ||||||||||  |||||    |||
Db        517 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 576

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627
              |    ||||  ||| |   | ||  | ||    ||  | || || || || |||| | ||
Db        577 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 636

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687
              |  | | |||   ||  | || || ||   | |     |||||    |  | |||||
Db        637 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 690

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747
              || |||||  | |   | || ||    |||||  |     || ||| | |   |   |
Db        691 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 750

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807
              |||||||| ||    |||||| || ||| |||| ||| | || ||||||      ||  ||
Db        751 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 810

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864
              |||  |   |  |||| ||||||||||||  |   | | | |      ||||||
Db        811 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 870

Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
              || || || |||||| |||| ||  | ||| | |||  |  ||||   || ||
Db        871 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 923


RESULT 12
ADJ98180
ID   ADJ98180 standard; DNA; 1188 BP.
XX
AC   ADJ98180;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Thale cress chlorophyllase 1 DNA.
XX
KW   phytol kinase; tocopherol biosynthesis; plant; drought resistance;
KW   thale cress; ds; chlorophyllase 1.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2004013312-A2.
XX
PD   12-FEB-2004.
XX
PF   05-AUG-2003; 2003WO-US025276.
XX
PR   05-AUG-2002; 2002US-0400689P.
PR   05-AUG-2003; 2003US-00634548.
XX
PA   (MONS ) MONSANTO TECHNOLOGY LLC.
XX
PI   Norris SR,  Lincoln K,  Abad MS,  Eilers R,  Hartsuyker KK;
PI   Hirschberg J,  Karunanandaa B,  Moshiri F,  Stein JC,  Valentin HE;
PI   Venkatesh TV;
XX
DR   WPI; 2004-157125/15.
XX
PT   New phytol kinase polynucleotides, useful in mediating tocopherol
PT   biosynthesis and in producing plants with increased drought resistance.
XX
PS   Claim 23; SEQ ID NO 18; 189pp; English.
XX
CC   The invention relates to a novel substantially purified nucleic acid
CC   molecule encoding a phytol kinase. The nucleic acid molecules and
CC   polypeptides of the invention may be useful in mediating tocopherol
CC   biosynthesis and in producing plants with increased drought resistance.
CC   The current sequence is that of the thale cress chlorophyllase 1 DNA of
CC   the invention.
XX
SQ   Sequence 1188 BP; 327 A; 255 C; 281 G; 325 T; 0 U; 0 Other;
```

```
  Query Match             14.4%;  Score 163.4;  DB 12;  Length 1188;
  Best Local Similarity   54.3%;  Pred. No. 1.1e-31;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;

Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210
Db        157 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 216

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270
Db        217 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 276

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330
Db        277 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 336

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390
Db        337 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 396

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450
Db        397 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 456

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507
Db        457 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 516

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567
Db        517 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 576

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627
Db        577 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 636

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687
Db        637 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 690

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747
Db        691 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 750

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807
Db        751 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 810

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864
Db        811 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGCAAAGAAAAAAGTCTGAGATGAGG 870

Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
Db        871 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 923
```

```
RESULT 13
AAD38077
ID   AAD38077 standard; cDNA; 1174 BP.
XX
AC   AAD38077;
XX
DT   10-SEP-2002  (first entry)
XX
DE   Soybean chlorophyllase cDNA from sfl1.pk0046.f8 clone.
XX
KW   Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW   enzyme; gene; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             12. .959
FT                   /*tag=  a
FT                   /product= "Soybean chlorophyllase protein from
FT                   sfl1.pk0046.f8 clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23779.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 52; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is soybean
CC   chlorophyllase cDNA from sfl1.pk0046.f8 clone
XX
SQ   Sequence 1174 BP; 318 A; 226 C; 260 G; 370 T; 0 U; 0 Other;

  Query Match             13.3%;  Score 151;  DB 6;  Length 1174;
  Best Local Similarity   55.0%;  Pred. No. 1.8e-28;
  Matches  433;  Conservative    0;  Mismatches  330;  Indels   24;  Gaps    6;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
Db         97 CATCCAATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTG 156

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
Db        157 GCTCATACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTG 216

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
Db        217 AGCTCCTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGA 276

Qy        323 TCGCCGGACC------AGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATT 376
Db        277 GTGTACGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATT 336

Qy        377 GGTTATC---AGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCA 433
Db        337 GGCTAGCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATA 396

Qy        434 AATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGA 493
Db        397 AATTGGTTTTATCAGGTCACAGCAAGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTT 456

Qy        494 AATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATG 553
Db        457 ATGCT------AAAACTAACCTCAAGTTTTCAGCACTAGTAGGCATAGACCCTGTGGCTG 510

Qy        554 GAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTG 613
```

```
Db        511 GCCCATGTAAATCTTGCGAAACATTTCCTCCTATTCTCACTGGCATGTCCCAATCCTTCA 570

Qy        614 ACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACC 673

Db        571 ATTTGAACATA---CCCATTGTTGTAATTGGCACTGGGCTAGGCCCAGAGAAGGCTAATT 627

Qy        674 CATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTC 733

Db        628 TTTTTATTCCACCATGTGCTCCTGATGGGGTGAACCATAAGGAGTTTTTCAATAAGTGCA 687

Qy        734 AAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATG 793

Db        688 AACCCCCTTGTGCACATTTTGTTGCAACTGAGTATGGTCACATGGACATGTTGGATGATG 747

Qy        794 ATACAAAAGGGATTAGAGGG---AAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAG--- 847

Db        748 TGACACCTGGCTTAATTGGGTCAATATTGTCAAATTGTATATGCAAGGATGGGAAGGGTC 807

Qy        848 AGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATT 907

Db        808 CTAGGGACTTGATGAGAAGGACCGTGGGAGGGTTGGTTGTGGCCTTCTTAAGGGCACAGT 867

Qy        908 TGGAAGG 914

Db        868 TGAATGG 874
```

RESULT 14
AAD38082
ID AAD38082 standard; cDNA; 1242 BP.
XX
AC AAD38082;
XX
DT 10-SEP-2002 (first entry)
XX
DE Wheat chlorophyllase cDNA from wle1n.pk0058.a4:fis clone.
XX
KW Wheat; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
OS Triticum aestivum.
XX

FH	Key	Location/Qualifiers
FT	CDS	47. .1006
FT		/*tag= a
FT		/product= "Wheat chlorophyllase protein from
FT		wle1n.pk0058.a4:fis clone"
FT		/EC_number= "3.1.1.14"

XX
PN WO200229022-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031059.
XX
PR 05-OCT-2000; 2000US-0238161P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon EB, Cahoon RE, Thorpe CJ;
XX
DR WPI; 2002-444102/47.
DR P-PSDB; AAE23784.
XX
PT An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT used to produce transformed plants that have controlled induction or
PT postponement of senescence.
XX
PS Claim 5; Page 59; 69pp; English.
XX
CC The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC in chlorophyll degradation in plant cell senescence. Sequences of the
CC invention are used to produce a plant by transforming a plant cell with
CC chlorophyllase and regenerating a plant from the transformed plant cell.
CC They may also be used to transform cells. The plants that are produced
CC can have economical importance as they can allow for the controlled
CC induction or postponement of senescence. The present sequence is wheat
CC chlorophyllase cDNA from wle1n.pk0058.a4:fis clone
XX
SQ Sequence 1242 BP; 266 A; 382 C; 343 G; 251 T; 0 U; 0 Other;

Query Match 13.1%; Score 148.2; DB 6; Length 1242;
Best Local Similarity 55.2%; Pred. No. 9.7e-28;
Matches 357; Conservative 0; Mismatches 278; Indels 12; Gaps 3;

```
Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Db        173 GCGGCGCCGACACCACCGATCCCGGTGCTGATCGTCGCACCCAAGGATGCAGGAACCTAC 232

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Db        233 CCCGTGGCCATGCTCTTGCACGGCTTCTTCCTCCATAACCACTTCTACGAACACCTTCTC 292

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTT---ATATAGTATCGCC 327

Db        293 CGGCACGTCGCATCCCACGGCTTCATCATTGTCGCGCCCCAGTTCAGCATCAGTATCATA 352

Qy        328 GGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTA 387

Db        353 CCTTCGGGTGACGCAGAGGACATCGCCGCGGCAGCCAAGGTGGCAGACTGGCTCCCCGAC 412

Qy        388 GGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCC 447

Db        413 GGCCTCCCGTCCGTGCTGCCCAAAGGCGTCGAGCCGGAGCTCTCGAAGCTCGCCTTGGCC 472

Qy        448 GGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCC 507

Db        473 GGCCACAGCCGAGGAGGCCACACGGCTTTCTCCCTGGCCTTGGGGCACGCCAAGACC--- 529

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Db        530 ---CAGCTAACCTTCTCCGCGCTCATCGGACTCGACCCCGTCGCCGGCACGGGGAAGTCC 586

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Db        587 TCCCAGCTCCAGCCCAAGATCCTCACCTACGAGCCGTCCTCCTTCGGCATGG---CGATG 643

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Db        644 CCGGTGCTGGTCATCGGCACCGGGCTCGGCGAGGAGAAGAAGAACATATTCTTCCCTCCC 703

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Db        704 TGCGCACCCAAGGACGTGAACCACGCCGAGTTCTACCGCGAGTGCAGGCCGCCCTGCTAC 763

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGA 794

Db        764 TACTTTGTGACCAAGGACTACGGGCATCTGGACATGCTGGACGACGA 810
```

RESULT 15
AAD38078
ID AAD38078 standard; cDNA; 1104 BP.
XX
AC AAD38078;
XX
DT 10-SEP-2002 (first entry)
XX
DE Soybean chlorophyllase cDNA from sfl1n1.pk002.m10:fis clone.
XX
KW Soybean; chlorophyllase; chlorophyll degradation; plant cell senescence;
KW enzyme; gene; ss.
XX
OS Glycine max.

```
XX
FH   Key             Location/Qualifiers
FT   CDS             26. .967
FT                   /*tag= a
FT                   /product= "Soybean chlorophyllase protein from
FT                   sfl1n1.pk002.m10:fis clone"
FT                   /EC_number= "3.1.1.14"
XX
PN   WO200229022-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-US031059.
XX
PR   05-OCT-2000; 2000US-0238161P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon EB,  Cahoon RE,  Thorpe CJ;
XX
DR   WPI; 2002-444102/47.
DR   P-PSDB; AAE23780.
XX
PT   An isolated polynucleotide encoding a plant chlorophyllase polypeptide,
PT   used to produce transformed plants that have controlled induction or
PT   postponement of senescence.
XX
PS   Claim 5; Page 53; 69pp; English.
XX
CC   The invention relates to plant chlorophyllase (EC 3.1.1.14) proteins and
CC   polynucleotides encoding such proteins. Chlorophyllase enzyme is involved
CC   in chlorophyll degradation in plant cell senescence. Sequences of the
CC   invention are used to produce a plant by transforming a plant cell with
CC   chlorophyllase and regenerating a plant from the transformed plant cell.
CC   They may also be used to transform cells. The plants that are produced
CC   can have economical importance as they can allow for the controlled
CC   induction or postponement of senescence. The present sequence is soybean
CC   chlorophyllase cDNA from sfl1n1.pk002.m10:fis clone
XX
SQ   Sequence 1104 BP; 308 A; 225 C; 249 G; 322 T; 0 U; 0 Other;

  Query Match             12.8%;  Score 145.4;  DB 6;  Length 1104;
  Best Local Similarity   53.8%;  Pred. No. 4.9e-27;
  Matches  421;  Conservative    0;  Mismatches  341;  Indels   21;  Gaps    5;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
              | || || |||||    || || ||||| |   || | |   ||| ||    |    |
Db        108 CATCCAGTGCTTCCTTCTCACCTCCAAAACCATTGCTAATTTTTACACCAACTGTTCCTG 167

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
              | |  || || ||  |  || || |||||||  |   |  |   ||| | || ||| ||
Db        168 GCGCATACCCTGTAATATTGTTCGTCCATGGCTTTTTCATTCGCAATTTCTACTACTCAA 227

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
              ||||  |    ||  |    || ||||| |||||  || | ||||||||  | | |   |
Db        228 AGCTCCTAGCCCACATAGTCTCACATGGATTCATAATCGTTGCTCCTCAACTGTTTTCCA 287

Qy        323 TCGCCGGACC---AGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGT 379
              |     ||     |  |     ||| |  |||   | |   |  ||  |||||||
Db        288 ATGGGCTTCCTATGTATGGACCCACTGAAGTGGAATATGCAGGAAAAGTTGCGGATTGGA 347

Qy        380 TATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTG 439
              || | |  |  ||| | || || ||||||| | | ||   | | ||  |   |||| | |
Db        348 TAGCTGAGGAGCTTCAACATTTGCTTCCAGAGAACGTTGAAGCCAATTTGGACAAACTGG 407

Qy        440 CCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG 499
              ||  || || || ||  | ||||| ||||| |  |||||  || ||  |          |
Db        408 TTCTATCAGGTCACAGTAGGGGTGGGAAAACTGTATTTGCTGTGGCTCT------TGGTC 461

Qy        500 GGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAG 559
                  | |||||  ||| | ||  || |  | || |||||  || || | ||| |||
Db        462 ATGCAAAAACTAATCTCAAGTTTTCAGCACTTGTAGGCATAGACCCTGTGGCTGGCACAT 521

Qy        560 GGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAG 619
              |||          ||| | |||||   |  |  ||  |   |||   || |||||  |
Db        522 CTAAATATTGTAGAACACGTCCTCATATTCTCACTGGCAAGCCACGGTCCTTTGATTT-- 579

Qy        620 ACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTAT 679
              ||||  |||  |    || || ||  | ||  | ||   |       |  ||   |
Db        580 -GAAAATGCCAGTTGAAGTAATTGGCACTGGATTGGGCCCAGAGAAGCTTAATTGTTGTA 638

Qy        680 TCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTC 739
              ||||||||||| |||   || |||||  |    ||||||||    || ||  ||   |
Db        639 CTCCACCGTGTGCTCCTGATGGGGTGAACTATAAGGAGTTCTTCAACGAGTGCAAACCCC 698

Qy        740 CAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAA 799
              |       | || ||||     | ||||| ||  |||||||| |  ||||||| |||
Db        699 CTTGTGCTAAATTTGTTGTAGCAAAGTATGGTCACATGGACATGTTGAATGATGACACAC 758

Qy        800 AAGGGATTAGAGGGA---AGAGTTCTTATTGTTTGTGTAAGAATGG------TGAAGAGA 850
              ||||  |  ||||     |   ||  | ||| ||||||||||||||      |    |
Db        759 CAGGGCTAATTGGGACATTGGTGTCAAAGTGTATGTGTAAGAATGGGACGACGGGTCCTA 818

Qy        851 GGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGG 910
              ||      ||||| ||    |  ||| ||  | ||||| || || |||| |||  | |||
Db        819 GGGACTTGATGAGAAGGACCACTGGAGGGTTGGTTGTGTCCTTCTTGAGGGCACAATTGA 878

Qy        911 AAG 913
              | |
Db        879 ATG 881
```

Search completed: March 20, 2006, 14:14:24
Job time : 671.906 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:06:13 ; Search time 1051.94 Seconds
(without alignments)
8922.325 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

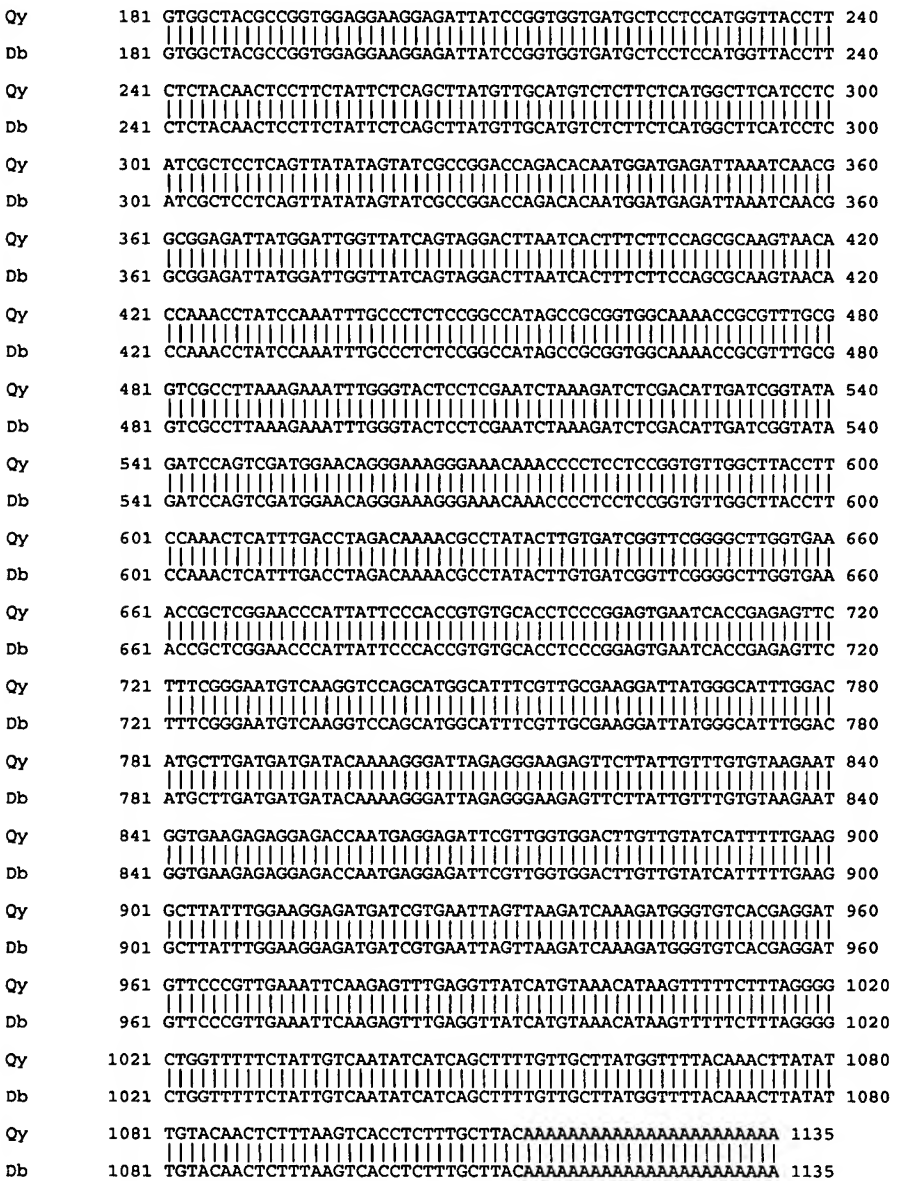
Result No.	Score	% Query Match	Length	DB	ID	Description
1	1135	100.0	1135	7	US-10-634-548-19	Sequence 19, Appl
2	367.4	32.4	1216	9	US-10-381-123-3	Sequence 3, Appli
3	256.8	22.6	1444	9	US-10-381-123-13	Sequence 13, Appl
4	237.4	20.9	987	9	US-10-381-123-1	Sequence 1, Appli
5	201.6	17.8	1302	9	US-10-381-123-5	Sequence 5, Appli
6	201.6	17.8	2439	8	US-10-425-115-140359	Sequence 140359,
7	185.6	16.4	796	7	US-10-424-599-86252	Sequence 86252, A
8	184	16.2	1125	7	US-10-425-114-26342	Sequence 26342, A
9	179.4	15.8	635	7	US-10-424-599-131410	Sequence 131410,
10	163.4	14.4	975	3	US-09-938-842A-158	Sequence 158, App
11	163.4	14.4	975	3	US-09-938-842A-158	Sequence 158, App
12	163.4	14.4	1188	7	US-10-634-548-18	Sequence 18, Appl
13	151	13.3	1174	9	US-10-381-123-7	Sequence 7, Appli
14	148.2	13.1	1242	9	US-10-381-123-17	Sequence 17, Appl
15	145.4	12.8	1104	9	US-10-381-123-9	Sequence 9, Appli
16	142.8	12.6	1125	9	US-10-381-123-11	Sequence 11, Appl
17	135.4	11.9	1307	7	US-10-424-599-45715	Sequence 45715, A
18	128.2	11.3	1155	7	US-10-437-963-94232	Sequence 94232, A
19	116.2	10.2	751	8	US-10-425-115-70835	Sequence 70835, A
20	116	10.2	1244	7	US-10-425-114-25386	Sequence 25386, A
21	116	10.2	1244	7	US-10-425-114-25387	Sequence 25387, A
22	91.4	8.1	1068	7	US-10-424-599-34940	Sequence 34940, A
23	88.4	7.8	1156	9	US-10-381-123-25	Sequence 25, Appl
24	85.2	7.5	538	7	US-10-424-599-34941	Sequence 34941, A
25	80.8	7.1	512	8	US-10-425-115-41107	Sequence 41107, A
26	71.4	6.3	916	9	US-10-381-123-15	Sequence 15, Appl
27	68.8	6.1	1004	7	US-10-424-599-68342	Sequence 68342, A
28	67.2	5.9	708	7	US-10-437-963-34072	Sequence 34072, A
29	52.2	4.6	1236	8	US-10-425-115-166593	Sequence 166593,
30	52.2	4.6	1274	9	US-10-381-123-27	Sequence 27, Appl
31	49	4.3	516	7	US-10-021-323-6457	Sequence 6457, Ap
c 32	48	4.2	487	7	US-10-021-323-2841	Sequence 2841, Ap
33	47.4	4.2	467	7	US-10-021-323-6541	Sequence 6541, Ap
c 34	47.4	4.2	552	7	US-10-021-323-93	Sequence 93, Appl
35	46.8	4.1	579	8	US-10-357-930-23827	Sequence 23827, A
36	46.8	4.1	579	8	US-10-357-930-29732	Sequence 29732, A
37	46.4	4.1	10039	6	US-10-311-455-2016	Sequence 2016, Ap
c 38	45.8	4.0	419	3	US-09-960-352-11234	Sequence 11234, A
c 39	45.8	4.0	629	6	US-10-125-968-78	Sequence 78, Appl
c 40	45.8	4.0	3673778	6	US-10-312-841-1	Sequence 1, Appli
41	45.6	4.0	728	7	US-10-767-701-3821	Sequence 3821, Ap
c 42	45.4	4.0	391	8	US-10-357-930-8140	Sequence 8140, Ap
43	45.4	4.0	524	7	US-10-767-701-3830	Sequence 3830, Ap
c 44	45.4	4.0	626	8	US-10-357-930-60960	Sequence 60960, A
45	45.2	4.0	421	3	US-09-918-995-7003	Sequence 7003, Ap

ALIGNMENTS

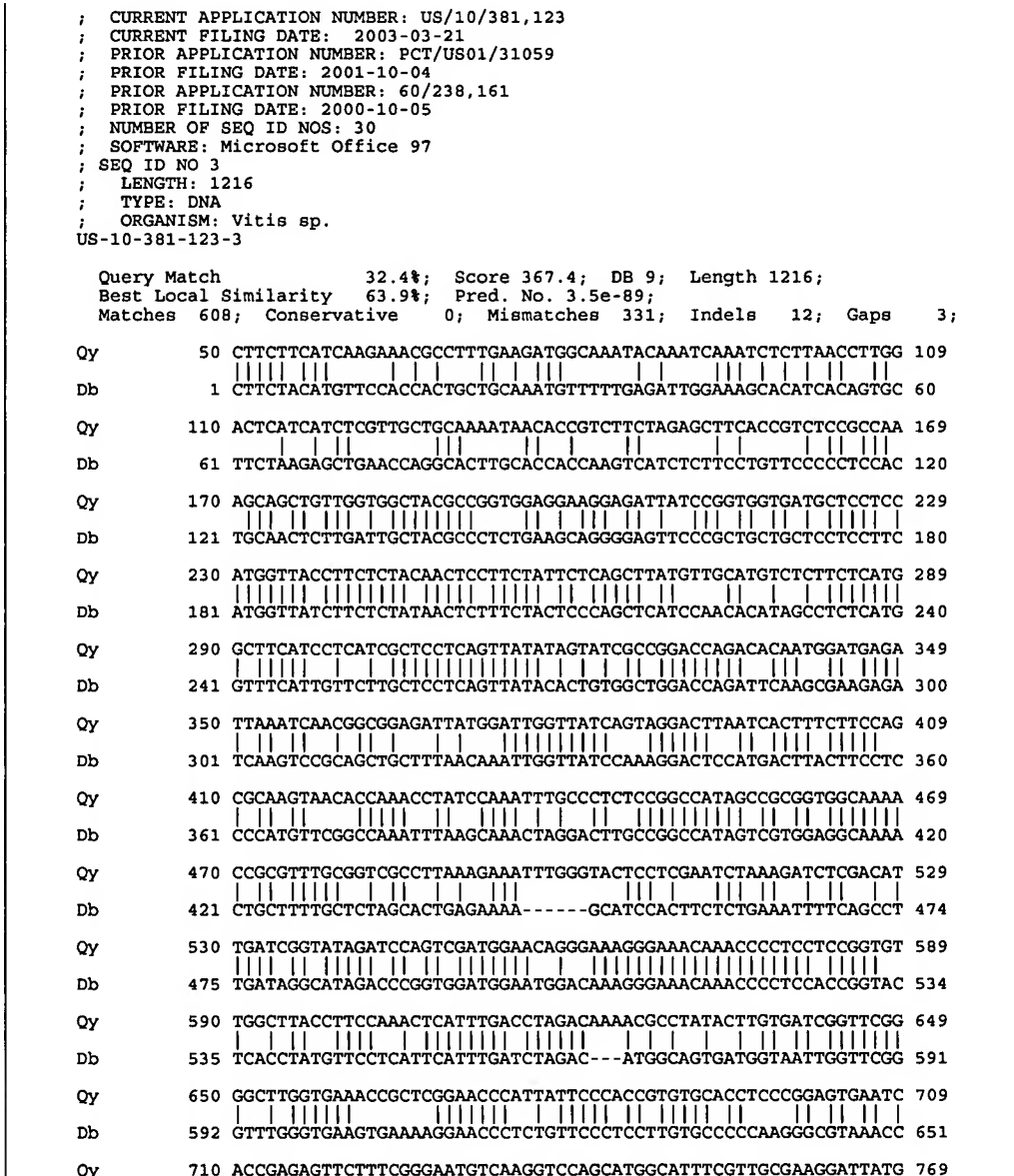
RESULT 1
US-10-634-548-19
; Sequence 19, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1ris, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen Kindle
; APPLICANT: Hirshberg, Joseph
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentin, Henry E.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
; FILE REFERENCE: Ren-01-125
; CURRENT APPLICATION NUMBER: US/10/634,548
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: us 60/400,689
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1135
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-634-548-19

Query Match 100.0%; Score 1135; DB 7; Length 1135;
Best Local Similarity 100.0%; Pred. No. 8.6e-299;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCA 60

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTG 180
```



RESULT 2
US-10-381-123-3
; Sequence 3, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT



; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1216
; TYPE: DNA
; ORGANISM: Vitis sp.
US-10-381-123-3

Query Match 32.4%; Score 367.4; DB 9; Length 1216;
Best Local Similarity 63.9%; Pred. No. 3.5e-89;
Matches 608; Conservative 0; Mismatches 331; Indels 12; Gaps 3;

```
Qy        890 CATTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTAAGATCAAAGATGGGT 949
Db        832 CATTTATGAAAGCTTATTTGGAAGGCGATAACAGCAGTCTAATCTCCATTAGAGATGGG- 890

Qy        950 GTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCATGTAAA 1000
Db        891 --CATGCTACTGCACCAGTGGAGCTTCAAAATGTTGAGTTTCTCGTGTGAA 939

RESULT 3
US-10-381-123-13
; Sequence 13, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
;   LENGTH: 1444
;   TYPE: DNA
;   ORGANISM: Gesneriana
US-10-381-123-13

  Query Match             22.6%;  Score 256.8;  DB 9;  Length 1444;
  Best Local Similarity   61.1%;  Pred. No. 6e-59;
  Matches  474;  Conservative    0;  Mismatches  287;  Indels   15;  Gaps    3;

Qy        154 TCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCG 213
Db        199 TCCGCACCACCTCGGAAGCCACCACTGATCGCAACTCCATTTGAGGAAGGCGAGTACCCA 258

Qy        214 GTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTG 273
Db        259 ACGCTTCTACTCCTTCATGGATTCATGCTTCACAACACTTTCTACTCTGAGCTTATCCAG 318

Qy        274 CATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCA 333
Db        319 CACATCGCATCCCATGGCTTTATTGTTGTTGTACCTCAGTTATATCTTGTAGCTACATGT 378

Qy        334 GACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTT 393
Db        379 GATAGTACGAATGGCATCAAGTCTGCTGCAAAAACAACAGATTGGTTGAAGGATGGACTG 438

Qy        394 AATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCAT 453
Db        439 CAAGATGTTCTCCCAACAAAAGTCAGACCAGACCTAAAGAAACTCGGACTGAGCGGACAT 498

Qy        454 AGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAAT 513
Db        499 AGCCGTGGCGGCAAAGATGCATTTGCTCTTGCACTAGGATATGCGAAG------ACTACA 552

Qy        514 CTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAA 573
Db        553 TTAAGCTTCTCAGCGCTGATCGGGATCGATCCTGTTGACGGAGTACGAAAGGGACACCAA 612

Qy        574 ACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATA 633
Db        613 ACCAATCCTCCTGTACTGAATTATATCCCTCACTCTTTGGAACT---CAAGATGCCGTCA 669

Qy        634 CTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCA 693
Db        670 TTAGTTATCGGAACAGGTTTAGGTGAATTGAAGAGGAACCTATTTGCC------TGTGCC 723

Qy        694 CCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTC 753
Db        724 CCTAAGGGAGTAAATCACCAAGATTTTTACGATGAGTGTTCTTCCCCGTCTTGCCATTTT 783

Qy        754 GTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGG 813
Db        784 GTTGCTAAGGACTATGGTCACATGGACATGCTAGATGATCATACTGACGGACCTAGAGGT 843

Qy        814 AAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGATTCGTT 873
Db        844 ATTGTGTCATACTGCATGTGTCGAAATGGAGAATCAAGAACACCCATGCGACTGTTTGTA 903

Qy        874 GGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATT 929
Db        904 GGTGGGATCGTGGTTGCTTTTCTGAAAGCCTGCTTACAAGGTGATAAACAGGATTT 959

RESULT 4
US-10-381-123-1
; Sequence 1, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
;   LENGTH: 987
;   TYPE: DNA
;   ORGANISM: Vitis sp.
US-10-381-123-1

  Query Match             20.9%;  Score 237.4;  DB 9;  Length 987;
  Best Local Similarity   56.3%;  Pred. No. 9.6e-54;
  Matches  551;  Conservative    0;  Mismatches  411;  Indels   17;  Gaps    5;

Qy        158 CGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCCGGTGG 217
Db         23 CTTCCCCTCCCAAGCCATTGTTGATTGTTACACCAACCATTCAAGGGACATACCCAGTTC 82

Qy        218 TGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTTGCATG 277
Db         83 TCTTGTTTCTTCATGGCTTCGAGCTCCGCAACACCTTCTACACTCAGCTCCTTCAACTCA 142

Qy        278 TCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACCAGACA 337
Db        143 TTTCTTCCCATGGATTCATTGTGGTGGCTCCTCAGTTATACGGACTATTACCTCCTTCTG 202

Qy        338 CAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACTTAATC 397
Db        203 GAATTCAAGAGATCAAATCAGCAGCAGCAGTCACAAATTGGCTATCCTCAGGCCTTCAAT 262

Qy        398 ACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCATAGCC 457
Db        263 CTGTGCTCCCAGAAAATGTGAAACCAGACCTACTCAAGCTTGCTCTTTCAGGCCACAGCA 322

Qy        458 GCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAATCTAA 517
Db        323 GAGGGGGAAAGACAGCATTTGCTCTGGCACTAGGGTATGCTGATACATCC------CTCA 376

Qy        518 AGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGAAACAAACCC 577
Db        377 ACTTCTCAGCCCTACTAGGACTAGACCCTGTTGGTGGGTTGAGTAAATGTTGCCAAACAG 436

Qy        578 CTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGCCTATACTTG 637
Db        437 TTCCCAAAATCCTAACCTATGTTCCTCATTCCTTCAATCTAG---CAATCCCAGTTTGCG 493
```

```
Qy        638 TGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGTGTGCACCTC 697
Db        494 TAATCGGCACGGGGTTGGGCGATGAGCCAAGGAACTGCCTAACATGTCCATGTGCCCCAG 553

Qy        698 CCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGCATTTCGTTG 757
Db        554 ATGGAGTGAACCATGTAGAGTTTTTCAGTGAGTGTAAACCTCCTTGTTCCCACTTTGTGA 613

Qy        758 CGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTAGAGGGAAGA 817
Db        614 CTACTGAATATGGTCACTTGGACATGTTAGATGATCATCTTTCAGGCTGCATTGGGGCGA 673

Qy        818 GTTCTTATTGTTTGTGTAAGAATGGTGAAG---AGAGGAGACCAATGAGGAGATTCGTTG 874
Db        674 TTTCGGGTTATATCTGCAAGAGTGGGAAGGGTCCTAGGGACCCCATGAGGAGATGTGTGG 733

Qy        875 GTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAATTAGTTA 934
Db        734 GTGGCCTTTTTGTTGCATTCTTGAAGGCTTATTTGGAAGGTCAGACTGGAGATT---TCA 790

Qy        935 AGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTTTGAGGTTATCA 994
Db        791 AAGCCATTGTTGATGAACCTGATCTGGCTCCTGTGAAGCTTGATCCTGTTGAGTTCATAG 850

Qy        995 TGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTG 1054
Db        851 AAGCATGAAATCATGGATTCTTTGTAATGCTATAAGGAGTGTGA--CAGCAGATTTTTTT 908

Qy       1055 TTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAA 1114
Db        909 TTCGGTTCGTCTTTGGGAACAAAGATTCAAATAAACGAAACTGTGTTGTATTGCTAAAAA 968

Qy       1115 AAAAAAAAAAAAAAAAAAA 1133
Db        969 AAAAAAAAAAAAAAAAAAA 987
```

RESULT 5
US-10-381-123-5
; Sequence 5, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Zea mays
US-10-381-123-5

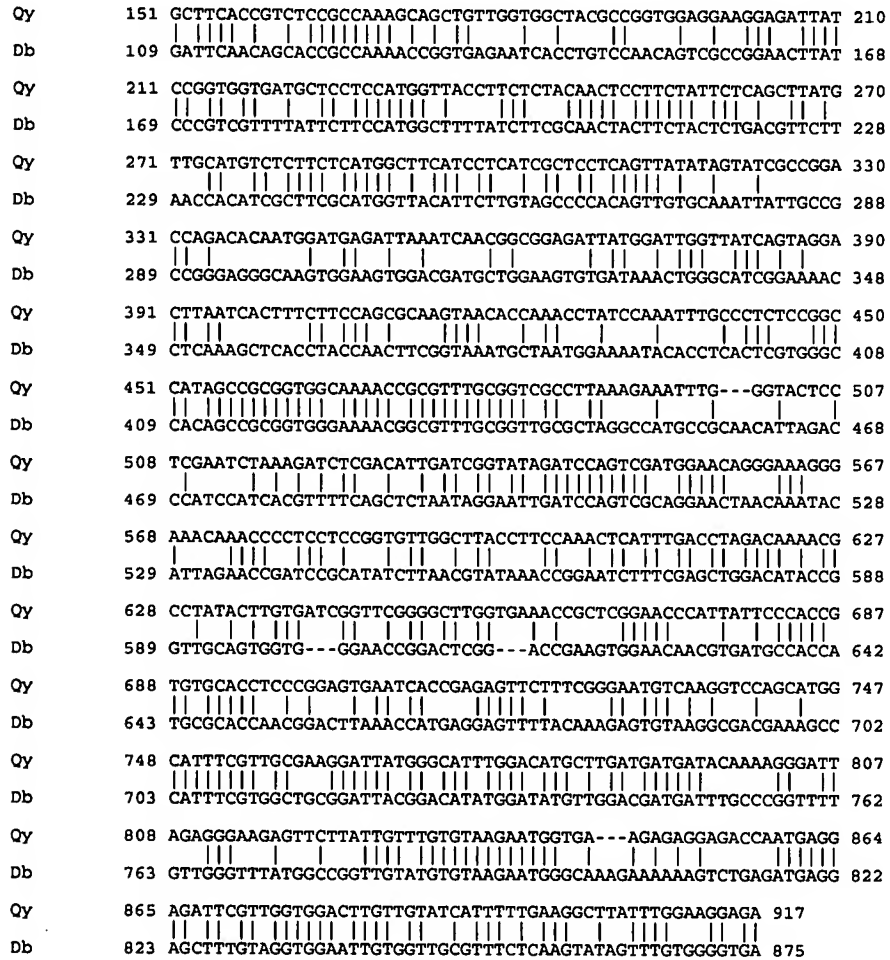
```
  Query Match             17.8%;  Score 201.6;  DB 9;  Length 1302;
  Best Local Similarity   54.6%;  Pred. No. 6.7e-44;
  Matches  449;  Conservative    0;  Mismatches  364;  Indels    9;  Gaps    2;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
Db        241 CGTCCAGCGCGGCCAGCCTGCCGCCGAAGCCGCTGCTCGTCGCCGCGCCGCGCGAGACTG 300

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
Db        301 GGGAGTACCCGGTGATCCTGTTCCTACACGGCTACCTCGCCGTCAACTCCTTCTACTCCC 360

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
Db        361 AGCTGTTCGAGCACGTCGCCTCCCATGGCTTTATCGTCGTCGGACCTCAGCTGTACACCA 420

Qy        323 TCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTAT 382
Db        421 TATCTGGGGCCGACACCACCGAGGAGATCAACTCAGCGGCGGCCGTCATCGACTGGCTAG 480

Qy        383 CAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCC 442
Db        481 CCACCGGGCTGCCGTCAACTCTGCCACTCGGCGTCCGCGCGAACCTAACCAAGGTGTCCA 540

Qy        443 TCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGT 502
Db        541 TCTCCGGCCACAGTCGCGGCGGGAAGGTGGCGTTCGCGCTGGCGTTGGCCACGCCAAGG 600

Qy        503 ACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGA 562
Db        601 CCAAGCTCGCTGTCCCTCTCGCCGCCGTCGTCGCCGTGGACCCGGTGGACGGCATGGGCG 660

Qy        563 AAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACA 622
Db        661 TGGGCAAGCAGACACCGCCGCCGATCCTCACGGGCAGGCACGGCTCGCTGCACGTGG--- 717

Qy        623 AAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCC 682
Db        718 GTGCCCCCGCCATGGTCATCGGCACGGGGCTCGGCGAGCTGCCCCGCGGCTCACTGCTCC 777

Qy        683 CACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCT------TTCGGGAATGTCAAG 736
Db        778 CGCCGTGCGCGCCCCGGGGCGTCAGTCACGCGGCCTTCTACGACGAGCTGGACGGCGCGG 837

Qy        737 GTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATA 796
Db        838 CGCCAGCGTGCCACCTGGTGGTCAGGGACTACGGGCACACGGACATGATGGACGACGACA 897

Qy        797 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 856
Db        898 CGCCGGGCGCCAGGGGGATGCTCACGCGCACCATCTGCAGGAGCGGCGGGGCCAGGCGC 957

Qy        857 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 916
Db        958 CCATGCGCCGCTTCGTGGCCGGCGCCACCGTCGCGTTCCTCAAGAAATGGGTGGCTGGGG 1017

Qy        917 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGG 958
Db       1018 ACGCCGCGGCGATGGACAGCATCACGGCGCGGCCGGACCAGG 1059
```

RESULT 6
US-10-425-115-140359
; Sequence 140359, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 140359
; LENGTH: 2439
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59490C.1
US-10-425-115-140359

```
  Query Match             17.8%;  Score 201.6;  DB 8;  Length 2439;
  Best Local Similarity   54.6%;  Pred. No. 9.6e-44;
```

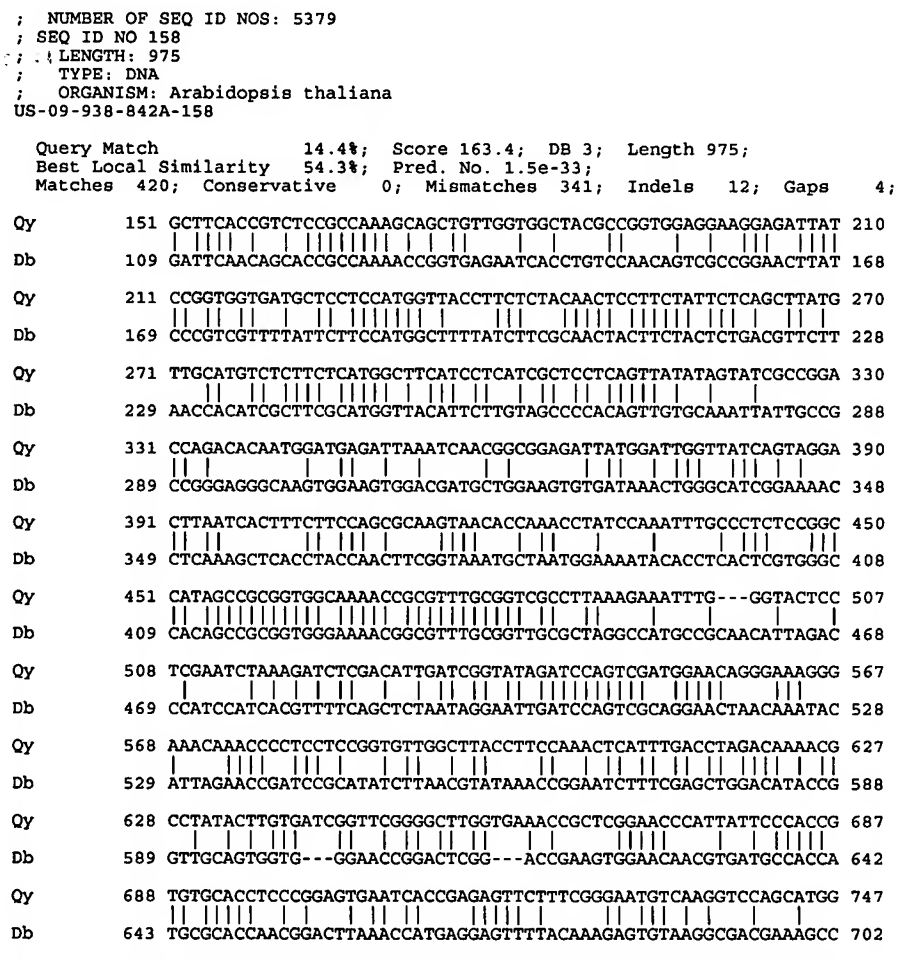
```
  Matches  449;  Conservative    0;  Mismatches  364;  Indels    9;  Gaps    2;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
              | || || ||  |       ||||| |||| |||| | || ||  |||||   |||  |
Db        282 CGTCCAGCGCGGCCAGCCTGCCGCCGAAGCCGCTGCTCGTCGCCGCGCCGCGCGAGACTG 341

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
              | || || |||||| |  || |||| || || ||||  |  ||||||||||| || |
Db        342 GGGAGTACCCGGTGATCCTGTTCCTACACGGCTACCTCGCCGTCAACTCCTTCTACTCCC 401

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
              ||||  |   ||| ||| | || |||||||| ||| || |||  |||||| | || | |
Db        402 AGCTGTTCGAGCACGTCGCCTCCCATGGCTTTATCGTCGTCGGACCTCAGCTGTACACCA 461

Qy        323 TCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTAT 382
              |  | ||  | |||||  |  || ||||| || ||| ||||||   | || || ||| ||
Db        462 TATCTGGGGCCGACACCACCGAGGAGATCAACTCAGCGGCGGCCGTCATCGACTGGCTAG 521

Qy        383 CAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCC 442
              |     || ||          ||| |||      ||      | |||||| |||| |  ||
Db        522 CCACCGGGCTGCCGTCAACTCTGCCACTCGGCGTCCGCGCGAACCTAACCAAGGTGTCCA 581

Qy        443 TCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGT 502
              |||||||||| || ||||| || ||     ||||| ||| | || ||    |       |
Db        582 TCTCCGGCCACAGTCGCGGCGGGAAGGTGGCGTTCGCGCTGGCGTTGGGCCACGCCAAGG 641

Qy        503 ACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGA 562
              |       | |    || | |  | |||   | || || || || || | |  ||
Db        642 CCAAGCTCGCTGTCCCTCTCGCCGCCGTCGTCGCCGTGGACCCGGTGGACGGCATGGGCG 701

Qy        563 AAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACA 622
              || || || || || || ||| |  |  |  |  |   ||| |  || | | |
Db        702 TGGGCAAGCAGACACCGCCGCCGATCCTCACGGGCAGGCACGGCTCGCTGCACGTGG--- 758

Qy        623 AAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCC 682
              | ||     | || |||||  ||||||| || ||     | ||   | || |  |||
Db        759 GTGCCCCCGCCATGGTCATCGGCACGGGGCTCGGCGAGCTGCCCCGCGGCTCACTGCTCC 818

Qy        683 CACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGG----- 737
              | ||||| || || |  || || | ||||   |  ||||     ||    || ||
Db        819 CGCCGTGCGCGCCCCGGGGCGTCAGTCACGCGGCCTTCTACGACGAGCTGCACGGCGCGG 878

Qy        738 -TCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATA 796
               ||||| || ||  | || |  | ||| || |||||  |||||| | || || ||| |
Db        879 CACCAGCGTGCCACCTGGTGGTCAGGGACTACGGGCACACGGACATGATGGACGACGACA 938

Qy        797 CAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGAC 856
              |   ||    || |||| |    |      | || | ||  || |  | |||   |
Db        939 CGCCGGGCGCCAGGGGGATGCTCACGCGCACCATCTGCAGGAGCGGCGGGGCCAGGGCGC 998

Qy        857 CAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAG 916
              | ||| |  | |||||  |  ||        ||  | ||  | |||   |  |||  || |
Db        999 CCATGCGCCGCTTCGTGGCCGGCGCCACCGTCGCGTTCCTCAAGAAATGGGTGGCTGGGG 1058

Qy        917 ATGATCGTGAATTAGTTAAGATCAAAGATGGGTGTCACGAGG 958
              | |      |  | |  |  ||||  |   ||    || |||
Db       1059 ACGCCGCGGCGATGGACAGCATCACGGCGCGGCCGGACCAGG 1100


RESULT 7
US-10-424-599-86252
; Sequence 86252, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 86252
;   LENGTH: 796
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_48899C.1
US-10-424-599-86252

  Query Match             16.4%;  Score 185.6;  DB 7;  Length 796;
  Best Local Similarity   65.5%;  Pred. No. 1.2e-39;
  Matches  288;  Conservative    0;  Mismatches  149;  Indels    3;  Gaps    1;

Qy        509 CGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGGA 568
              | |||||| ||| | ||  | || || ||  | |||||||| |||||||  |  |||||
Db          8 CTAATCTCAAGTTTTCAGCCTTAATAGGAGTGGATCCAGTTGATGGAATGGACAAAGGAA 67

Qy        569 AACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACGC 628
              | |||||||| || || ||   | | ||  |||| ||||||||||  | |   | | |
Db         68 AGCAAACCCCCCCACCTGTTCTCACCTATGTTCCTAACTCATTTGATTTTG---ATATGG 124

Qy        629 CTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCGT 688
              | |  | || || ||||| || || ||||||       |||| || || || || ||
Db        125 CAGTGATGGTCATAGGTTCAGGTCTAGGTGAAGTGAAAAGGAATCCTTTGTTTCCCCCAC 184

Qy        689 GTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGGC 748
              | | | |   || || || ||   | | |||||    || ||| |    || || |||
Db        185 GCGTCCGTAAGGGTGTCAACCATGAAAACTTCTTCAATGAGTGTAAGAAACCTGCTTGGT 244

Qy        749 ATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATTA 808
              |||| || || |||||||||||| || |  |||||||| |||||||| || || || ||||
Db        245 ATTTTGTGGCAAAGGATTATGGCCACTCTGACATGCTAGATGATGACACCAAGGGAATTA 304

Qy        809 GAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAGGAGAT 868
              | ||||||   | |  | ||| | || || ||||| ||    |||| ||| ||||| || |
Db        305 GGGGGAAAGCTACCAACTGTCTATGCAAAAATGGAGAGTCAAGGAAACCTATGAGAAGGT 364

Qy        869 TCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCGTGAAT 928
              | ||||| ||| |  ||||  |||| ||||| ||||| ||| | |  ||| ||   || |
Db        365 TTGTTGGAGGAGTCATTGTTGCATTCTTGAAAGCTTACTTGCATGATGATAATGAGGACT 424

Qy        929 TAGTTAAGATCAAAGATGGG 948
              | | |  || | |||  ||
Db        425 TGTTGACCATAAGAGACAGG 444


RESULT 8
US-10-425-114-26342
; Sequence 26342, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26342
;   LENGTH: 1125
;   TYPE: DNA
;   ORGANISM: Brassica napus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB4156-010-D3_FLI
US-10-425-114-26342

  Query Match             16.2%;  Score 184;  DB 7;  Length 1125;
  Best Local Similarity   56.1%;  Pred. No. 3.9e-39;
```

```
  Matches  431;  Conservative    0;  Mismatches  325;  Indels   12;  Gaps    4;

Qy        139 CCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAG 198
              |||||  | |  | ||    |  | || |||||||    |  |  |  |  | ||    |
Db        135 CCGTCGGCAACCGATTTGACCGCACCTCCAAAGCCTGTAATAATCACCTCCCCAACCGTC 194

Qy        199 GAAGGAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTAT 258
              |  |||  ||| || || ||   |  || |||||||  | |  |||    |||| ||||||
Db        195 GCCGGAACTTACCCCGTCGTCTTATTCTTCCATGGATTCTATCTTCGTAACTACTTCTAC 254

Qy        259 TCTCAGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATAT 318
              |||  |  ||||    || ||  ||||||||||||| |||   |  | || || ||| | |
Db        255 TCTGATGTTATTAACCACGTAGCTTCTCATGGCTACATTGTTGTAGCCCCACAGCTGTGC 314

Qy        319 AGTATCGCCGGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGG 378
              |  ||       || |      |  ||  |  |      | |   |  | ||  | |||
Db        315 AAGATTTTGCCGCCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAAAGTGATAAACTGG 374

Qy        379 TTATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTT 438
              ||   |  || ||      || |||      |||||   | ||      | ||
Db        375 ACTTCGAAAAACCTCAAAGCTCACCTCCCAAGCTCAGTAAACGCTAATGGCAACTACACC 434

Qy        439 GCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTT 498
              || |||   ||||||||||||||||| ||||||||||||||||| || |||     |
Db        435 GCACTCGTGGGCCATAGCCGCGGTGGTAAAACCGCGTTTGCGGTTGCGTTAGGCCACGCC 494

Qy        499 G---GGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGA 555
              |      | |     | ||| | ||  |  |  | || ||||||||||| |  |||
Db        495 GCAACACTAGACCCATCCATCAAGTTTTCAGCTCTTGTAGGAATAGATCCAGTTGCAGGA 554

Qy        556 ACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGAC 615
              |   | ||| | | |  |||| |||     | ||  | |||   ||  | ||||| |||
Db        555 ATCAGCAAATGCATAAGAACCGATCCCGAAATCTTAACGTACAAACCGGAATCATTCGAC 614

Qy        616 CTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCA 675
              || ||||   ||  |  |   |||||||| |||| || ||   | |     | |||
Db        615 CTGGACATGCCGGTTGCA---GTGATCGGTACGGGTCTCGG---ACCGAAGAGTAACATG 668

Qy        676 TTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAA 735
              |  | |||| || ||||  | | |||||| ||    |||| | |   || ||| |
Db        669 CTGATGCCACCATGCGCACCAGCGGAAGTGAACCATGAGGAGTTTTATATTGAGTGTAAG 728

Qy        736 GGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGAT 795
              | | |     |||||||||| ||   |||||| || ||| |||| ||| | || ||| ||
Db        729 GCTACCAAGGGGCATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATAAT 788

Qy        796 ACAAAAGGGATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAA---GAGAGG 852
              |  ||   |||   |   |   |||| ||||||||||  || ||    | | |
Db        789 TTGCCCGGTTTTGTCGGGTTTATGGCGGGTTGTATGTGTAAGAACAGTAAACGCAAAAAG 848

Qy        853 AGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAG 900
              ||    ||||| || || ||||||||| |||| ||  | ||| | |||
Db        849 AGTGAGATGAGAAGCTTTGTTGGTGGAATTGTGGTTGCGTTTCTAAAG 896


RESULT 9
US-10-424-599-131410
; Sequence 131410, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131410
;   LENGTH: 635
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(635)
;   OTHER INFORMATION: unsure at all n locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_89671C.1
US-10-424-599-131410

  Query Match           15.8%;  Score 179.4;  DB 7;  Length 635;
  Best Local Similarity   66.3%;  Pred. No. 5.1e-38;
  Matches  273;  Conservative    0;  Mismatches  137;  Indels    2;  Gaps    1;

Qy        153 TTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTATCC 212
              ||  ||  | || ||||||   || ||| | || || ||  | || |  |||||||  ||
Db        226 TTTTCCACCACCACCAAAGTCTCTCTTGATTGCCACTCCTCTTGAAGGTGGAGATTTCCC 285

Qy        213 GGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATGTT 272
                |  || || | || ||||| || ||||| ||||| || || || || ||||| ||
Db        286 TCTTCTGCTGTTTCTGCATGGCTATCTTCTTTACAATTCATTTTACTCCCAGCTAATACA 345

Qy        273 GCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGACC 332
              |||||  |||||||||| |||||  |||| || |||||||| || |   | || |||||
Db        346 ACATGTTGCTTCTCATGGTTTCATTGTCATTGCCCCTCAGTTGTACACAGTGGCTGGACC 405

Qy        333 AGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGACT 392
              || ||||  |||||||| | ||  | || |  || |   | ||| | || | ||||||
Db        406 TGATACAAGTGATGAGATTCATTCTGCAGCTGCAATAACAAACTGGCTCTCTGAAGGACT 465

Qy        393 TAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGCCA 452
                | ||||||||| |  | |||| ||||||  |   ||| || || ||  | |||||
Db        466 CTGCAAGTTTCTTCCACCAAATGTAAGACCAAATTTGAGCAAGTTAGCACTTGCTGGCCA 525

Qy        453 TAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCCTCGAA 512
              ||| ||||| || ||||| |||||||| | ||  |||  || |     | |  ||  ||
Db        526 TAGTCGCGGAGGAAAAACAGCGTTTGCTCTTGCTCTAAGAAACTAAACATCCAACT--AA 583

Qy        513 TCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAA 564
              ||| ||| | ||  | || ||  |  | |||||||| |||||||  |  |||
Db        584 TCTCAAGTTTTCAGCCTTAATANGAGTGGATCCAGTTGATGGAATGGACAAA 635


RESULT 10
US-09-938-842A-158
; Sequence 158, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 158
;   LENGTH: 975
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

  Query Match           14.4%;  Score 163.4;  DB 3;  Length 975;
  Best Local Similarity   54.3%;  Pred. No. 1.5e-33;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;
```



RESULT 11
US-09-938-842A-158
; Sequence 158, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 158
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-158

Query Match 14.4%; Score 163.4; DB 3; Length 975;
Best Local Similarity 54.3%; Pred. No. 1.5e-33;
Matches 420; Conservative 0; Mismatches 341; Indels 12; Gaps 4;



RESULT 12
US-10-634-548-18
; Sequence 18, Application US/10634548
; Publication No. US20040045051A1
; GENERAL INFORMATION:
; APPLICANT: No. US20040045051A1ris, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert

```
;  APPLICANT:  Hartsuyker, Karen Kindle
;  APPLICANT:  Hirshberg, Joseph
;  APPLICANT:  Karunanandaa, Balasulojini
;  APPLICANT:  Moshiri, Farhad
;  APPLICANT:  Stein, Joshua C.
;  APPLICANT:  Valentin, Henry E.
;  APPLICANT:  Venkatesh, Tyamagondlu V.
;  TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
;  FILE REFERENCE: Ren-01-125
;  CURRENT APPLICATION NUMBER: US/10/634,548
;  CURRENT FILING DATE:  2003-08-05
;  PRIOR APPLICATION NUMBER: us 60/400,689
;  PRIOR FILING DATE: 2002-08-05
;  NUMBER OF SEQ ID NOS: 79
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
;   LENGTH: 1188
;   TYPE: DNA
;   ORGANISM: Synechocystis PCC6803
US-10-634-548-18

  Query Match             14.4%;  Score 163.4;  DB 7;  Length 1188;
  Best Local Similarity   54.3%;  Pred. No. 1.7e-33;
  Matches  420;  Conservative    0;  Mismatches  341;  Indels   12;  Gaps    4;

Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210
              | |||| |  | |||||||| | | ||    |  |    ||    |  |  |||  ||||
Db        157 GATTCAACAGCACCGCCAAAACCGGTGAGAATCACCTGTCCAACAGTCGCCGGAACTTAT 216

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270
              || || ||  |  || |||||||| |    |||   ||||| |||||| ||| |  || |
Db        217 CCCGTCGTTTTATTCTTCCATGGCTTTTATCTTCGCAACTACTTCTACTCTGACGTTCTT 276

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTATCGCCGGA 330
              ||  || |||| ||||| | ||| ||  | || || |||||  |   |
Db        277 AACCACATCGCTTCGCATGGTTACATTCTTGTAGCCCCACAGTTGTGCAAATTATTGCCG 336

Qy        331 CCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTAGGA 390
              || |       |  ||  |  |     | |      | ||  | |||  ||| | |
Db        337 CCGGGAGGGCAAGTGGAAGTGGACGATGCTGGAAGTGTGATAAACTGGGCATCGGAAAAC 396

Qy        391 CTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCCGGC 450
              || ||       || ||| |    ||||   | ||   |    |  | |||   |||
Db        397 CTCAAAGCTCACCTACCAACTTCGGTAAATGCTAATGGAAAATACACCTCACTCGTGGGC 456

Qy        451 CATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG---GGTACTCC 507
              || |||||||||| ||||| |||||||||| ||  ||    |    |    |     |
Db        457 CACAGCCGCGGTGGGAAAACGGCGTTTGCGGTTGCGCTAGGCCATGCCGCAACATTAGAC 516

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567
              |     | | | | ||  |  | || || || ||||||||||  |||||    |||
Db        517 CCATCCATCACGTTTTCAGCTCTAATAGGAATTGATCCAGTCGCAGGAACTAACAAATAC 576

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627
              |    ||||  ||| |   | ||  | ||    ||  | || || || || |||| | ||
Db        577 ATTAGAACCGATCCGCATATCTTAACGTATAAACCGGAATCTTTCGAGCTGGACATACCG 636

Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687
              |  | | |||   ||  | || || ||   | |     |||||    |  | |||||
Db        637 GTTGCAGTGGTG---GGAACCGGACTCGG---ACCGAAGTGGAACAACGTGATGCCACCA 690

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747
              || |||||  | |   | || ||    ||||| |      || ||| | |   |   |
Db        691 TGCGCACCAACGGACTTAAACCATGAGGAGTTTTACAAAGAGTGTAAGGCGACGAAAGCC 750

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAAAAGGGATT 807
              |||||||| ||   |||||| || ||| |||| ||| | || ||||||      ||  ||
Db        751 CATTTCGTGGCTGCGGATTACGGACATATGGATATGTTGGACGATGATTTGCCCGGTTTT 810

Qy        808 AGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGA---AGAGAGGAGACCAATGAGG 864
              |||  |   |  |||| ||||||||||||||  |   | | |      ||||||
Db        811 GTTGGGTTTATGGCCGGTTGTATGTGTAAGAATGGGCAAAGAAAAAAGTCTGAGATGAGG 870
```

```
Qy        865 AGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGA 917
              || || || |||||| |||| ||  | ||| | |||  |  ||||  || ||
Db        871 AGCTTTGTAGGTGGAATTGTGGTTGCGTTTCTCAAGTATAGTTTGTGGGGTGA 923


RESULT 13
US-10-381-123-7
; Sequence 7, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  TITLE OF INVENTION: Chlorophyllases
;  FILE REFERENCE: BB1477 PCT
;  CURRENT APPLICATION NUMBER: US/10/381,123
;  CURRENT FILING DATE:  2003-03-21
;  PRIOR APPLICATION NUMBER: PCT/US01/31059
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: 60/238,161
;  PRIOR FILING DATE: 2000-10-05
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 1174
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-381-123-7

  Query Match             13.3%;  Score 151;  DB 9;  Length 1174;
  Best Local Similarity   55.0%;  Pred. No. 4.1e-30;
  Matches  433;  Conservative    0;  Mismatches  330;  Indels   24;  Gaps    6;

Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202
              | || |  |||||  | || || ||||| |   ||||  |   ||| ||    | |   |
Db         97 CATCCAATGCTTCCTCCTCACCTCCAAAACCATTGTTAATCTTTACACCAACCGTGCCTG 156

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262
              |    || || ||  |  || ||  |||||| |     ||   |||   || ||| |||
Db        157 GCTCATACCCTGTAATATTGTTCTGCCATGGATTTTCCCTTCGCAATAGCTACTACTCTG 216

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322
              ||||  |    ||  |  |||| ||||| |||||  |  | |||||||||| | | | |
Db        217 AGCTCCTAGGCCACATAGCTTCACATGGATTCATAATTGTTGCTCCTCAGCTGTGTTGGA 276

Qy        323 TCGCCGGACC------AGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATT 376
              |   |  |       ||  |    |||||  ||||||   | | | |  || ||||||
Db        277 GTGTACGGTCTATGTTGGAACCTGGTGATGAAGTTAAATTTGCAGGGAAAGTTGTGGATT 336

Qy        377 GGTTATC---AGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCA 433
              || || |    |  || ||| | |   | |||||| | | ||   | | ||  |  |
Db        337 GGCTAGCCGAGGAGGGGCTTCAACCTCTGCTTCCAGAGAATGTTGAAGCCAAATTGGATA 396

Qy        434 AATTTGCCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGA 493
              |||| |   | || || || |||   |||||||||| |  ||||| || ||  |
Db        397 AATTGGTTTTATCAGGTCACAGCAAGGGTGGCAAAACTGTATTTGCTGTGGCACTTGGTT 456

Qy        494 AATTTGGGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATG 553
              |   |           | || || ||| | ||  || |  | || ||||| || || | ||
Db        457 ATGCT------AAAACTAACCTCAAGTTTTCAGCACTAGTAGGCATAGACCCTGTGGCTG 510

Qy        554 GAACAGGGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTG 613
              |  || | |||       ||||   ||||||  |  |  ||  | |  |  | || ||
Db        511 GCCCATGTAAATCTTGCGAAACATTTCCTCCTATTCTCACTGGCATGTCCCAATCCTTCA 570

Qy        614 ACCTAGACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACC 673
              |  |  ||| |   || ||  |||| || ||  | ||||| ||   |          ||
Db        571 ATTTGAACATA---CCCATTGTTGTAATTGGCACTGGGCTAGGCCCAGAGAAGGCTAATT 627

Qy        674 CATTATTCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTC 733
              ||  | ||||| ||||| |||   || ||||| ||    |||||  ||      | ||
Db        628 TTTTTATTCCACCATGTGCTCCTGATGGGGTGAACCATAAGGAGTTTTTCAATAAGTGCA 687

Qy        734 AAGGTCCAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATG 793
              ||   ||       ||||| ||||| |  || |||||  ||  |||||||| | |||||||
```

```
Db        688 AACCCCCTTGTGCACATTTTGTTGCAACTGAGTATGGTCACATGGACATGTTGGATGATG 747

Qy        794 ATACAAAAGGGATTAGAGGG---AAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAG--- 847

Db        748 TGACACCTGGCTTAATTGGGTCAATATTGTCAAATTGTATATGCAAGGATGGGAAGGGTC 807

Qy        848 AGAGGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATT 907

Db        808 CTAGGGACTTGATGAGAAGGACCGTGGGAGGGTTGGTTGTGGCCTTCTTAAGGGCACAGT 867

Qy        908 TGGAAGG 914

Db        868 TGAATGG 874
```

RESULT 14
US-10-381-123-17
; Sequence 17, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-381-123-17

Query Match 13.1%; Score 148.2; DB 9; Length 1242;
Best Local Similarity 55.2%; Pred. No. 2.5e-29;
Matches 357; Conservative 0; Mismatches 278; Indels 12; Gaps 3;

```
Qy        151 GCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAGGAGATTAT 210

Db        173 GCGGCGCCGACACCACCGATCCCGGTGCTGATCGTCGCACCCAAGGATGCAGGAACCTAC 232

Qy        211 CCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTCAGCTTATG 270

Db        233 CCCGTGGCCATGCTCTTGCACGGCTTCTTCCTCCATAACCACTTCTACGAACACCTTCTC 292

Qy        271 TTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTT---ATATAGTATCGCC 327

Db        293 CGGCACGTCGCATCCCACGGCTTCATCATTGTCGCGCCCCAGTTCAGCATCAGTATCATA 352

Qy        328 GGACCAGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGTTATCAGTA 387

Db        353 CCTTCGGGTGACGCAGAGGACATCGCCGCGGCAGCCAAGGTGGCAGACTGGCTCCCCGAC 412

Qy        388 GGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTGCCCTCTCC 447

Db        413 GGCCTCCCGTCCGTGCTGCCCAAAGGCGTCGAGCCGGAGCTCTCGAAGCTCGCCTTGGCC 472

Qy        448 GGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTGGGTACTCC 507

Db        473 GGCCACAGCCGAGGAGGCCACACGGCTTTCTCCCTGGCCTTGGGGCACGCCAAGACC--- 529

Qy        508 TCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAGGGAAAGGG 567

Db        530 ---CAGCTAACCTTCTCCGCGCTCATCGGACTCGACCCCGTCGCCGGCACGGGGAAGTCC 586

Qy        568 AAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAGACAAAACG 627

Db        587 TCCCAGCTCCAGCCCAAGATCCTCACCTACGAGCCGTCCTCCTTCGGCATGG---CGATG 643
```

```
Qy        628 CCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTATTCCCACCG 687

Db        644 CCGGTGCTGGTCATCGGCACCGGGCTCGGCGAGGAGAAGAAGAACATATTCTTCCCTCCC 703

Qy        688 TGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTCCAGCATGG 747

Db        704 TGCGCACCCAAGGACGTGAACCACGCCGAGTTCTACCGCGAGTGCAGGCCGCCCTGCTAC 763

Qy        748 CATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGA 794

Db        764 TACTTTGTGACCAAGGACTACGGGCATCTGGACATGCTGGACGACGA 810
```

RESULT 15
US-10-381-123-9
; Sequence 9, Application US/10381123
; Publication No. US20050081263A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours and Company
; TITLE OF INVENTION: Chlorophyllases
; FILE REFERENCE: BB1477 PCT
; CURRENT APPLICATION NUMBER: US/10/381,123
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/US01/31059
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,161
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Glycine max
US-10-381-123-9

Query Match 12.8%; Score 145.4; DB 9; Length 1104;
Best Local Similarity 53.8%; Pred. No. 1.4e-28;
Matches 421; Conservative 0; Mismatches 341; Indels 21; Gaps 5;

```
Qy        143 CTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTACGCCGGTGGAGGAAG 202

Db        108 CATCCAGTGCTTCCTTCTCACCTCCAAAACCATTGCTAATTTTTACACCAACTGTTCCTG 167

Qy        203 GAGATTATCCGGTGGTGATGCTCCTCCATGGTTACCTTCTCTACAACTCCTTCTATTCTC 262

Db        168 GCGCATACCCTGTAATATTGTTCGTCCATGGCTTTTTCATTCGCAATTTCTACTACTCAA 227

Qy        263 AGCTTATGTTGCATGTCTCTTCTCATGGCTTCATCCTCATCGCTCCTCAGTTATATAGTA 322

Db        228 AGCTCCTAGCCCACATAGTCTCACATGGATTCATAATCGTTGCTCCTCAACTGTTTTCCA 287

Qy        323 TCGCCGGACC---AGACACAATGGATGAGATTAAATCAACGGCGGAGATTATGGATTGGT 379

Db        288 ATGGGCTTCCTATGTATGGACCCACTGAAGTGGAATATGCAGGAAAAGTTGCGGATTGGA 347

Qy        380 TATCAGTAGGACTTAATCACTTTCTTCCAGCGCAAGTAACACCAAACCTATCCAAATTTG 439

Db        348 TAGCTGAGGAGCTTCAACATTTGCTTCCAGAGAACGTTGAAGCCAATTTGGACAAACTGG 407

Qy        440 CCCTCTCCGGCCATAGCCGCGGTGGCAAAACCGCGTTTGCGGTCGCCTTAAAGAAATTTG 499

Db        408 TTCTATCAGGTCACAGTAGGGGTGGGAAAACTGTATTTGCTGTGGCTCT------TGGTC 461

Qy        500 GGTACTCCTCGAATCTAAAGATCTCGACATTGATCGGTATAGATCCAGTCGATGGAACAG 559

Db        462 ATGCAAAAACTAATCTCAAGTTTTCAGCACTTGTAGGCATAGACCCTGTGGCTGGCACAT 521

Qy        560 GGAAAGGGAAACAAACCCCTCCTCCGGTGTTGGCTTACCTTCCAAACTCATTTGACCTAG 619

Db        522 CTAAATATTGTAGAACACGTCCTCATATTCTCACTGGCAAGCCACGGTCCTTTGATTT-- 579

Qy        620 ACAAAACGCCTATACTTGTGATCGGTTCGGGGCTTGGTGAAACCGCTCGGAACCCATTAT 679
```

```
Db        580 -GAAAATGCCAGTTGAAGTAATTGGCACTGGATTGGGCCCAGAGAAGCTTAATTGTTGTA 638

Qy        680 TCCCACCGTGTGCACCTCCCGGAGTGAATCACCGAGAGTTCTTTCGGGAATGTCAAGGTC 739
              ||||||||||| |||   || |||||  |    ||||||||    || ||  ||   |
Db        639 CTCCACCGTGTGCTCCTGATGGGGTGAACTATAAGGAGTTCTTCAACGAGTGCAAACCCC 698

Qy        740 CAGCATGGCATTTCGTTGCGAAGGATTATGGGCATTTGGACATGCTTGATGATGATACAA 799
              |         | || ||||       | ||||| ||  |||||||| |  ||||||| |||
Db        699 CTTGTGCTAAATTTGTTGTAGCAAAGTATGGTCACATGGACATGTTGAATGATGACACAC 758

Qy        800 AAGGGATTAGAGGGA---AGAGTTCTTATTGTTTGTGTAAGAATGG------TGAAGAGA 850
              |||| | |  ||||    |   ||  | ||| ||||||||||||      |     |
Db        759 CAGGGCTAATTGGGACATTGGTGTCAAAGTGTATGTGTAAGAATGGGACGACGGGTCCTA 818

Qy        851 GGAGACCAATGAGGAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGG 910
              ||      ||||| ||   |  ||| ||  | ||||| || || |||| |||  | |||
Db        819 GGGACTTGATGAGAAGGACCACTGGAGGGTTGGTTGTGTCCTTCTTGAGGGCACAATTGA 878

Qy        911 AAG 913
              | |
Db        879 ATG 881
```

Search completed: March 20, 2006, 14:42:22
Job time : 1056.94 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 14:07:47 ; Search time 453.902 Seconds
(without alignments)
5830.519 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq1:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq1:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq4:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43.2	3.8	627	6	US-09-925-065A-605854	Sequence 605854,
2	43	3.8	135462	7	US-10-330-773-37	Sequence 37, Appl
3	41.8	3.7	561	6	US-09-925-065A-343968	Sequence 343968,
4	40.8	3.6	561	6	US-09-925-065A-343967	Sequence 343967,
5	40.6	3.6	1721	8	US-10-947-249-91	Sequence 91, Appl
6	40.6	3.6	1721	12	US-11-122-329-76	Sequence 76, Appl
7	40.4	3.6	1842	9	US-11-096-568A-12837	Sequence 12837, A
8	40.2	3.5	652	6	US-09-925-065A-84331	Sequence 84331, A
9	40.2	3.5	652	6	US-09-925-065A-84332	Sequence 84332, A
10	40.2	3.5	2201	9	US-11-050-857-1056	Sequence 1056, Ap
11	40.2	3.5	2201	9	US-11-051-720-1711	Sequence 1711, Ap
12	40.2	3.5	2201	9	US-11-043-806-516	Sequence 516, App
13	40.2	3.5	2201	12	US-11-043-590-46	Sequence 46, Appl
14	40	3.5	498	6	US-09-925-065A-325092	Sequence 325092,
15	40	3.5	874	8	US-10-750-185-31978	Sequence 31978, A
16	40	3.5	874	8	US-10-750-623-31978	Sequence 31978, A
17	40	3.5	5160	12	US-11-108-172-690	Sequence 690, App
c 18	39.8	3.5	494	6	US-09-925-065A-566557	Sequence 566557,
c 19	39.8	3.5	520	6	US-09-925-065A-291203	Sequence 291203,
c 20	39.6	3.5	270	6	US-09-925-065A-202708	Sequence 202708,
c 21	39.6	3.5	1680	8	US-10-750-185-55400	Sequence 55400, A
c 22	39.6	3.5	1680	8	US-10-750-623-55400	Sequence 55400, A
23	39.6	3.5	9001	12	US-11-011-332A-4	Sequence 4, Appli
24	39.6	3.5	28536	12	US-11-011-332A-155	Sequence 155, App
25	39.4	3.5	551	6	US-09-925-065A-149499	Sequence 149499,
26	39.4	3.5	551	6	US-09-925-065A-149500	Sequence 149500,
c 27	39.4	3.5	552	6	US-09-925-065A-774116	Sequence 774116,
28	39.4	3.5	642	6	US-09-925-065A-714153	Sequence 714153,
29	39.4	3.5	2328	8	US-10-770-726-7	Sequence 7, Appli
30	39.4	3.5	2328	12	US-11-000-688-154	Sequence 154, App
c 31	39.4	3.5	9001	12	US-11-011-332A-4	Sequence 4, Appli
c 32	39.4	3.5	28536	12	US-11-011-332A-155	Sequence 155, App
33	39.2	3.5	615	6	US-09-925-065A-163852	Sequence 163852,
34	39.2	3.5	1662	12	US-11-028-922A-3	Sequence 3, Appli
c 35	39	3.4	560	6	US-09-925-065A-23149	Sequence 23149, A
36	38.8	3.4	629	6	US-09-925-065A-172188	Sequence 172188,
37	38.8	3.4	787	6	US-09-925-065A-77487	Sequence 77487, A
38	38.8	3.4	787	6	US-09-925-065A-77488	Sequence 77488, A
39	38.8	3.4	1076	6	US-09-925-065A-685537	Sequence 685537,
40	38.6	3.4	592	6	US-09-925-065A-843639	Sequence 843639,
41	38.6	3.4	2299	9	US-11-245-400-6	Sequence 6, Appli
42	38.6	3.4	138808	7	US-10-330-773-905	Sequence 905, App
c 43	38.6	3.4	1082144	12	US-11-117-187-211	Sequence 211, App
c 44	38.4	3.4	535	6	US-09-925-065A-269865	Sequence 269865,
c 45	38.4	3.4	3105	12	US-11-091-883-239	Sequence 239, App

ALIGNMENTS

RESULT 1
US-09-925-065A-605854
; Sequence 605854, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 605854
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-605854

Query Match 3.8%; Score 43.2; DB 6; Length 627;
Best Local Similarity 54.4%; Pred. No. 5.2;
Matches 87; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
Qy          1 AAAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAAATGTCCTCTTCTTCATCA 60
              || ||||  ||| ||||  |    || ||||||  ||| || | |     || ||    |
Db        235 AAGAAAAAGAAATAAAATCATGTATATTAAAGAGGAAACAATAAGCATCTTTTTTACATA 294

Qy         61 AGAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCT 120
              ||||| |  | |  || |   |||| ||||  |||| | |||| |||  |      ||
Db        295 AGAAATGATTGTCCAGGTCAAAAATTCAAAAGAAATTTTTTAAACTTCAAAAGTGAATGA 354

Qy        121 CGTTGCTGCAAAATAACACCGTCTTCTAGAGCTTCACCGT 160
                || |    |||| ||  | ||   ||| |   | | |
```

```
Db        355 GTATGTTTATCAATATCATAGGCTATAAGATCAGTATCCT 394


RESULT 2
US-10-330-773-37
; Sequence 37, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
;  APPLICANT: David W. Morris
;  APPLICANT:  Marc Malandro
;  TITLE OF INVENTION: Novel Compositions and Methods in Cancer
;  FILE REFERENCE: 529452001300
;  CURRENT APPLICATION NUMBER: US/10/330,773
;  CURRENT FILING DATE:  2002-12-27
;  NUMBER OF SEQ ID NOS: 981
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 135462
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(135462)
;   OTHER INFORMATION: n = A,T,C or G
US-10-330-773-37

  Query Match             3.8%;  Score 43;  DB 7;  Length 135462;
  Best Local Similarity   52.5%;  Pred. No. 31;
  Matches  117;  Conservative    0;  Mismatches  105;  Indels    1;  Gaps    1;

Qy          9 TAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCAAGAAACGC 68
              |||| ||||||||| |   |  ||| || |||||  | | ||| | |||   ||||
Db     124353 TAAAAAAAAGAAACAAAACAAAACAAAAAAAACCAGGACCTGTGCTTTCCCAAGTTCTTA 124412

Qy         69 CTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCTCGTTGCTG 128
              |||| |   |    |||  ||   |   ||   |   || || |  |  | | |  |||
Db     124413 CTTTAAGCTTTTAGAATGGAAGCAATTGCTTGCATAGTTAGAATGTTGGTGT-GGAGCTC 124471

Qy        129 CAAAATAACACCGTCTTCTAGAGCTTCACCGTCTCCGCCAAAGCAGCTGTTGGTGGCTAC 188
              |||||   |  |||   |   |  | | | | | | |||    |||||||||   |||
Db     124472 CAAAAGGCCCACGTTGGCAGTAAGTACTCTGACCCGGCCTCTACAGCTGTTGTCTGCTTT 124531

Qy        189 GCCGGTGGAGGAAGGAGATTATCCGGTGGTGATGCTCCTCCAT 231
              || || |   | ||   | ||| | |   ||| |   || |
Db     124532 CCCTGTCGTCCAGGGTTTTGTTCCCGAGTCAATGTTTGACCGT 124574


RESULT 3
US-09-925-065A-343968
; Sequence 343968, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 343968
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-343968

  Query Match             3.7%;  Score 41.8;  DB 6;  Length 561;
  Best Local Similarity   53.5%;  Pred. No. 10;
  Matches   85;  Conservative    1;  Mismatches   73;  Indels    0;  Gaps    0;

Qy        977 AAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGT 1036
              || |||  ||  | ||||  | || | |     |  : |     |   ||  | |||| |
Db        101 AATAGTGAGATATAATCAAATTAAGAGATTCAATGAYCTCATTACCAATTCATTTATTTT 160

Qy       1037 CAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAG 1096
                || ||  |||||||     | | | | || ||  |||||| || ||  | ||
Db        161 ACTATTCTTCCTCTTTTGTATGAAAAGATCTAACTAAAATATATTATAAAAAGCATTCTT 220

Qy       1097 TCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
                ||  |   || | |||  |||||||| || | || ||
Db        221 CCAAATTCCTGTTAACATTAAAAAAAAAGAAGAGAAGAA 259


RESULT 4
US-09-925-065A-343967
; Sequence 343967, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 343967
;   LENGTH: 561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-343967

  Query Match             3.6%;  Score 40.8;  DB 6;  Length 561;
  Best Local Similarity   52.4%;  Pred. No. 17;
  Matches   87;  Conservative    1;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        970 GAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTT 1029
              :||    || |||  ||  | ||||  | || | |     |    |     |   ||  |
Db         94 RAAGAAAAATAGTGAGATATAATCAAATTAAGAGATTCAATGACCTCATTACCAATTCAT 153

Qy       1030 CTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACT 1089
              ||||  |      || ||  |||||||     | | | | || ||  |||||| || ||
Db        154 TTATTTTACTATTCTTCCTCTTTTGTATGAAAAGATCTAACTAAAATATATTATAAAAAG 213

Qy       1090 CTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              | ||    ||  |   || | |||  |||||||| || | || ||
Db        214 CATTCTTCCAAATTCCTGTTAACATTAAAAAAAAAGAAGAGAAGAA 259


RESULT 5
US-10-947-249-91
; Sequence 91, Application US/10947249
; Publication No. US20050287541A1
; GENERAL INFORMATION:
;  APPLICANT: Akira NAKAGAWARA
;  APPLICANT:  Miki OHIRA
```

```
;  APPLICANT:  Shin ISHII
;  APPLICANT:  Takeshi GOTO
;  APPLICANT:  Hiroyuki KUBO
;  APPLICANT:  Takahiro HIRATA
;  APPLICANT:  Yasuko YOSHIDA
;  APPLICANT:  Saichi YAMADA
;  TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma and Meth
;  TITLE OF INVENTION:  Predicting the Prognosis of Neuroblastoma
;  FILE REFERENCE: 117007
;  CURRENT APPLICATION NUMBER: US/10/947,249
;  CURRENT FILING DATE:  2004-09-23
;  PRIOR APPLICATION NUMBER: US 60/505,614
;  PRIOR APPLICATION NUMBER: 2003-09-25
;  NUMBER OF SEQ ID NOS: 200
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
;   LENGTH: 1721
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-947-249-91

  Query Match             3.6%;  Score 40.6;  DB 8;  Length 1721;
  Best Local Similarity   58.8%;  Pred. No. 27;
  Matches   70;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1017 GGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTT 1076
              ||||  || ||   | |||||         |||||| ||| ||| |  ||  |||
Db       1584 GGGGAAGGATTGGGTCTTGTCCCCCAACACAGCTTCTGTGGCTGACTGTAATACTGTACA 1643

Qy       1077 ATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              |   | |   ||  || ||  |   | |   || | ||||||||||||||||||||||||
Db       1644 ACTGTTTCTGACCATTAAATGCTGTTGTACTCTGAAAAAAAAAAAAAAAAAAAAAAAAAA 1702


RESULT 6
US-11-122-329-76
; Sequence 76, Application US/11122329
; Publication No. US20060019272A1
; GENERAL INFORMATION:
;  APPLICANT: Geraci, Mark
;  APPLICANT:  Bull, Todd
;  APPLICANT:  Voelkel, Norbert
;  APPLICANT:  Coldren, Chris
;  TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
;  TITLE OF INVENTION:  Expression Analysis of Peripheral Blood Cells
;  FILE REFERENCE: 2848-54
;  CURRENT APPLICATION NUMBER: US/11/122,329
;  CURRENT FILING DATE:  2005-05-03
;  PRIOR APPLICATION NUMBER: 60/568,129
;  PRIOR FILING DATE: 2004-05-03
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76
;   LENGTH: 1721
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-122-329-76

  Query Match             3.6%;  Score 40.6;  DB 12;  Length 1721;
  Best Local Similarity   58.8%;  Pred. No. 27;
  Matches   70;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1017 GGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTT 1076
              ||||  || ||   | |||||         |||||| ||| ||| |  ||  |||
Db       1584 GGGGAAGGATTGGGTCTTGTCCCCCAACACAGCTTCTGTGGCTGACTGTAATACTGTACA 1643

Qy       1077 ATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              |   | |   ||  || ||  |   | |   || | ||||||||||||||||||||||||
Db       1644 ACTGTTTCTGACCATTAAATGCTGTTGTACTCTGAAAAAAAAAAAAAAAAAAAAAAAAAA 1702


RESULT 7
US-11-096-568A-12837
; Sequence 12837, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12837
;   LENGTH: 1842
;   TYPE: DNA
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1842)
;   OTHER INFORMATION: Ceres Seq. ID no. 14303706
US-11-096-568A-12837

  Query Match             3.6%;  Score 40.4;  DB 9;  Length 1842;
  Best Local Similarity   65.6%;  Pred. No. 30;
  Matches   59;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       1046 CAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTT 1105
              |||  | | ||   ||||  ||    | | |||||||||      ||||||     | |||
Db       1702 CAGTATATCTTTGATATGTATTGTATATCATATATTGTATTGAACTTTAATAGTTCACTT 1761

Qy       1106 TGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              ||  |   |||||||||||||||||||||||
Db       1762 TGTGTTTGAAAAAAAAAAAAAAAAAAAAAA 1791


RESULT 8
US-09-925-065A-84331
; Sequence 84331, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84331
;   LENGTH: 652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-84331

  Query Match             3.5%;  Score 40.2;  DB 6;  Length 652;
  Best Local Similarity   84.9%;  Pred. No. 24;
  Matches   45;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1083 TACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
              || |  | || ||||||||||||  |||| ||||||||||||||||||||||||
Db        132 TATATTTGTTAAAGTCACCTCTTCACTTAAAAAAAAAAAAAAAAAAAAAAAAAA 184


RESULT 9
US-09-925-065A-84332
```

```
; Sequence 84332, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84332
;   LENGTH: 652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-84332

  Query Match             3.5%;  Score 40.2;  DB 6;  Length 652;
  Best Local Similarity   84.9%;  Pred. No. 24;
  Matches   45;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1083 TACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAAA 1135
             || |  | || ||||||||||||  |||| |||||||||||||||||||||||||
Db        132 TATATTTGTTAAAGTCACCTCTTCACTTAAAAAAAAAAAAAAAAAAAAAAAAAA 184


RESULT 10
US-11-050-857-1056
; Sequence 1056, Application US/11050857
; Publication No. US20060040278A1
; GENERAL INFORMATION:
;  APPLICANT: Compugen Ltd
;  TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
;  TITLE OF INVENTION:  THEREOF FOR DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 1847.1005
;  CURRENT APPLICATION NUMBER: US/11/050,857
;  CURRENT FILING DATE:  2005-01-27
;  NUMBER OF SEQ ID NOS: 1150
; SEQ ID NO 1056
;   LENGTH: 2201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-050-857-1056

  Query Match             3.5%;  Score 40.2;  DB 9;  Length 2201;
  Best Local Similarity   70.1%;  Pred. No. 35;
  Matches   54;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       1059 TTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAA 1118
             ||| ||||| |  || ||  |||| ||    |||| | | |  |  |||| | |  ||||
Db       2122 TTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCCAAAA 2181

Qy       1119 AAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||
Db       2182 AAAAAAAAAAAAAAAAA 2198


RESULT 11
US-11-051-720-1711
; Sequence 1711, Application US/11051720
; Publication No. US20060046257A1
; GENERAL INFORMATION:
;  APPLICANT: Compugen Ltd
;  TITLE OF INVENTION: NOVEL NUCLEOTIDE AND AMINO ACID SEQUENCES, AND ASSAYS AND METHODS
;  TITLE OF INVENTION:  THEREOF FOR DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 1847.1002
;  CURRENT APPLICATION NUMBER: US/11/051,720
;  CURRENT FILING DATE:  2005-01-27
;  NUMBER OF SEQ ID NOS: 1780
; SEQ ID NO 1711
;   LENGTH: 2201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-051-720-1711

  Query Match             3.5%;  Score 40.2;  DB 9;  Length 2201;
  Best Local Similarity   70.1%;  Pred. No. 35;
  Matches   54;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       1059 TTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAA 1118
             ||| ||||| |  || ||  |||| ||    |||| | | |  |  |||| | |  ||||
Db       2122 TTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCCAAAA 2181

Qy       1119 AAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||
Db       2182 AAAAAAAAAAAAAAAAA 2198


RESULT 12
US-11-043-806-516
; Sequence 516, Application US/11043806
; Publication No. US20060051774A1
; GENERAL INFORMATION:
;  APPLICANT: Compugen Ltd
;  TITLE OF INVENTION: Novel Nucleotide and Amino Acid Sequences, and Assays and Methods
;  TITLE OF INVENTION:  thereof for Diagnosis of Prostate Cancer
;  FILE REFERENCE: 1847.1003
;  CURRENT APPLICATION NUMBER: US/11/043,806
;  CURRENT FILING DATE:  2005-01-27
;  NUMBER OF SEQ ID NOS: 575
; SEQ ID NO 516
;   LENGTH: 2201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-043-806-516

  Query Match             3.5%;  Score 40.2;  DB 9;  Length 2201;
  Best Local Similarity   70.1%;  Pred. No. 35;
  Matches   54;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       1059 TTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAA 1118
             ||| ||||| |  || ||  |||| ||    |||| | | |  |  |||| | |  ||||
Db       2122 TTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCCAAAA 2181

Qy       1119 AAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||
Db       2182 AAAAAAAAAAAAAAAAA 2198


RESULT 13
US-11-043-590-46
; Sequence 46, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
;  APPLICANT: Compugen Ltd
;  TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use theree
;  FILE REFERENCE: 1847.1011
;  CURRENT APPLICATION NUMBER: US/11/043,590
;  CURRENT FILING DATE:  2005-01-27
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 46
;   LENGTH: 2201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-043-590-46
```

```
  Query Match             3.5%;  Score 40.2;  DB 12;  Length 2201;
  Best Local Similarity   70.1%;  Pred. No. 35;
  Matches   54;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy       1059 TTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAA 1118
             ||| ||||| |  || ||  |||| ||    |||| | | |  | |||| | |  ||||
Db       2122 TTAAGGTTTCAACAAATTTCATTGCACTTTCCTTTCAATAAAGTTGTTGCATTCCCAAAA 2181

Qy       1119 AAAAAAAAAAAAAAAAA 1135
              |||||||||||||||||
Db       2182 AAAAAAAAAAAAAAAAA 2198


RESULT 14
US-09-925-065A-325092
; Sequence 325092, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single
;  TITLE OF INVENTION:  Nucleotide Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.135
;  CURRENT APPLICATION NUMBER: US/09/925,065A
;  CURRENT FILING DATE:  2001-08-08
;  PRIOR APPLICATION NUMBER: US 60/243,096
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 60/252,147
;  PRIOR FILING DATE: 2000-11-20
;  PRIOR APPLICATION NUMBER: US 60/250,092
;  PRIOR FILING DATE: 2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/261,766
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/289,846
;  PRIOR FILING DATE: 2001-05-09
;  NUMBER OF SEQ ID NOS: 957086
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 325092
;   LENGTH: 498
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-065A-325092

  Query Match             3.5%;  Score 40;  DB 6;  Length 498;
  Best Local Similarity   55.1%;  Pred. No. 25;
  Matches   76;  Conservative    1;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          2 AAAAAAGTAAAGAAAAGAAAAACTAATAAAGAACAAAAAAAATGTCCTCTTCTTCATCAA 61
              ||||||  ||| |||| |||||  || || || |||||  ||  |  |    |  ||||
Db        230 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAGCAAAAGACCTGAACAGTCACTTCTCAA 289

Qy         62 GAAACGCCTTTGAAGATGGCAAATACAAATCAAATCTCTTAACCTTGGACTCATCATCTC 121
               |:|   | |  |||  | |||  |  |   |||         | |   ||| |||||
Db        290 AARAAAACATACAAGTGGCCAAGAAACATATAAAAAATACTCCATATCACTAATCATTGG 349

Qy        122 GTTGCTGCAAAATAACAC 139
               |   ||||||  || ||
Db        350 ATAAGTGCAAATCAAAAC 367


RESULT 15
US-10-750-185-31978
; Sequence 31978, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;  APPLICANT: MMI GENOMICS, INC.
;  APPLICANT:  DeNISE, Sue K.
;  APPLICANT:  KERR, Richard
;  APPLICANT:  ROSENFELD, David
;  APPLICANT:  HOLM, Tom
;  APPLICANT:  BATES, Stephen
;  APPLICANT:  FANTIN, Dennis
;  TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;  FILE REFERENCE: MMI1100-2
;  CURRENT APPLICATION NUMBER: US/10/750,185
;  CURRENT FILING DATE:  2003-12-31
;  PRIOR APPLICATION NUMBER: US 60/437,482
;  PRIOR FILING DATE: 2002-12-31
;  NUMBER OF SEQ ID NOS: 64922
;  SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31978
;   LENGTH: 874
;   TYPE: DNA
;   ORGANISM: Bovine  19866881260462
US-10-750-185-31978

  Query Match             3.5%;  Score 40;  DB 8;  Length 874;
  Best Local Similarity   45.7%;  Pred. No. 29;
  Matches  139;  Conservative    0;  Mismatches  165;  Indels    0;  Gaps    0;

Qy        804 GATTAGAGGGAAGAGTTCTTATTGTTTGTGTAAGAATGGTGAAGAGAGGAGACCAATGAG 863
              ||  |  |||  | || |  |||| | | |      ||  |  || ||  |  |||
Db        542 GACAAATGGGGCTATTTTTGCTTGTCTTTCTTTTGGTGCAGTGCAGTGGCTAAGAATACA 601

Qy        864 GAGATTCGTTGGTGGACTTGTTGTATCATTTTTGAAGGCTTATTTGGAAGGAGATGATCG 923
                | |    ||  || ||| | |||   |||||       | |          || |
Db        602 AGTTCTGGACACTGTCCTCGTTCTGTCAGACTTGAAAAAAATTATTTTTTTCCCGGAGCC 661

Qy        924 TGAATTAGTTAAGATCAAAGATGGGTGTCACGAGGATGTTCCCGTTGAAATTCAAGAGTT 983
              | ||||  || |  |  || ||||||  | |||         | |  || ||   | |
Db        662 TTAATTTCTTCATCTGTAACATGGGTTAAATGAGAGATAATAAGATTGAACTCCGAATTA 721

Qy        984 TGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATATC 1043
              ||||| | | |   ||  || |   ||           |  | | | | |||    ||
Db        722 TGAGGATTTAAGTCAACGATGACCCTTGGGGGTCTAACATCTGTGTTTCCTGTGTTGCTC 781

Qy       1044 ATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTC 1103
              | ||| ||  ||||||||||   || |  | | | |         |||  ||    | | |
Db        782 ACCAGATTCAGTTGCTTATGAGCTTGCCCATTCAGAAGTCTTTCCTCAATAGAAAAACCC 841

Qy       1104 TTTG 1107
              || |
Db        842 TTAG 845


Search completed: March 20, 2006, 14:23:27
Job time : 454.902 secs
```

THIS PAGE LEFT BLANK

OM nucleic - nucleic search, using sw model

Run on: March 20, 2006, 13:57:59 ; Search time 226.218 Seconds
(without alignments)
8918.522 Million cell updates/sec

Title: US-10-634-548-19
Perfect score: 1135
Sequence: 1 aaaaaaagtaaagaaaagaa.....aaaaaaaaaaaaaaaaaaa 1135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	44.2	3.9	1979	3	US-09-461-325-30	Sequence 30, Appl
	2	44.2	3.9	1979	3	US-10-012-542-30	Sequence 30, Appl
	3	44.2	3.9	1979	3	US-10-115-123-30	Sequence 30, Appl
	4	43	3.8	198	2	US-08-330-108-16	Sequence 16, Appl
	5	43	3.8	198	6	PCT-US92-10087-16	Sequence 16, Appl
c	6	42.8	3.8	1637	3	US-09-270-767-11283	Sequence 11283, A
	7	42.4	3.7	991	3	US-08-924-747-25	Sequence 25, Appl
	8	42.4	3.7	991	3	US-09-247-373B-25	Sequence 25, Appl
	9	42.4	3.7	991	3	US-09-296-715-25	Sequence 25, Appl
	10	42	3.7	3265	3	US-09-832-129-13	Sequence 13, Appl
c	11	41.6	3.7	640681	3	US-09-790-988-1	Sequence 1, Appli
	12	40.4	3.6	2186	3	US-10-033-301-23	Sequence 23, Appl
	13	40	3.5	1278	3	US-09-008-697A-11	Sequence 11, Appl
	14	40	3.5	2862	2	US-08-148-209A-1	Sequence 1, Appli
	15	39.6	3.5	1380	3	US-09-881-556A-1	Sequence 1, Appli
	16	39.6	3.5	7260	3	US-09-917-254-31	Sequence 31, Appl
	17	39.6	3.5	7260	3	US-09-919-497-24	Sequence 24, Appl
	18	39.4	3.5	1722	3	US-09-134-001C-1976	Sequence 1976, Ap
	19	39.4	3.5	2007	3	US-08-747-221B-36	Sequence 36, Appl
c	20	39.4	3.5	2007	3	US-08-747-221B-38	Sequence 38, Appl
	21	39.4	3.5	2007	3	US-09-005-051-36	Sequence 36, Appl
c	22	39.4	3.5	2007	3	US-09-005-051-38	Sequence 38, Appl
	23	39.4	3.5	2007	3	US-09-403-942F-36	Sequence 36, Appl
c	24	39.4	3.5	2007	3	US-09-403-942F-38	Sequence 38, Appl
	25	39.2	3.5	1141	3	US-09-806-708B-22	Sequence 22, Appl
	26	39.2	3.5	1582	3	US-08-545-196B-10	Sequence 10, Appl
	27	39.2	3.5	1582	3	US-08-545-196B-12	Sequence 12, Appl
	28	39.2	3.5	1737	2	US-08-202-056-4	Sequence 4, Appli
	29	39.2	3.5	1737	2	US-08-076-093A-3	Sequence 3, Appli
	30	39.2	3.5	1737	2	US-08-701-265-3	Sequence 3, Appli
	31	39.2	3.5	1737	2	US-08-284-586-3	Sequence 3, Appli
	32	39.2	3.5	1737	2	US-08-805-478-3	Sequence 3, Appli
	33	39.2	3.5	1737	2	US-08-802-627A-3	Sequence 3, Appli
	34	39.2	3.5	1737	2	US-08-801-238-3	Sequence 3, Appli
	35	39.2	3.5	1737	2	US-08-801-228-3	Sequence 3, Appli
	36	39.2	3.5	1737	3	US-09-104-296-3	Sequence 3, Appli
	37	39.2	3.5	1737	6	PCT-US94-06380-2	Sequence 2, Appli
	38	39.2	3.5	4709	3	US-08-956-171E-587	Sequence 587, App
	39	39.2	3.5	4709	3	US-08-781-986A-587	Sequence 587, App
c	40	39.2	3.5	77586	3	US-09-949-016-13220	Sequence 13220, A
c	41	39.2	3.5	77586	3	US-09-949-016-13221	Sequence 13221, A
c	42	39	3.4	1664976	3	US-08-916-421B-1	Sequence 1, Appli
c	43	39	3.4	1664976	3	US-09-692-570-1	Sequence 1, Appli
	44	38.8	3.4	213456	3	US-09-820-007-3	Sequence 3, Appli
	45	38.6	3.4	1939	3	US-09-719-083A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-461-325-30
; Sequence 30, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (968)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-461-325-30

Query Match 3.9%; Score 44.2; DB 3; Length 1979;
Best Local Similarity 53.8%; Pred. No. 0.084;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```
Qy        967 GTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTT 1026
              ||| ||| || ||   |||   | ||||| |||   ||  | | |||  || | |
Db       1802 GTTCAAAGTCTAGCTATTGGTATAATCATCTAATATTATATATATCTCCAGTGCCCCTGA 1861

Qy       1027 TTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACA 1086
              || |||  |  ||  |   |  ||| |     |||  | ||  |     |||    | |
Db       1862 ATTTTATGTTTGATGACTATATATTTGGGCATATATCTTGTTGGATTAGAATAAATAAAA 1921
```

```
Qy       1087 ACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                ||  | |  | |     || | ||||||||||||||||||||||||
Db       1922 CACTTTATATTTTCATGAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAA 1970


RESULT 2
US-10-012-542-30
; Sequence 30, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029P1
;  CURRENT APPLICATION NUMBER: US/10/012,542
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 1979
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (968)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-012-542-30

  Query Match             3.9%;  Score 44.2;  DB 3;  Length 1979;
  Best Local Similarity   53.8%;  Pred. No. 0.084;
  Matches   91;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        967 GTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTT 1026
              ||| ||| || ||   |||   | ||||| |||   ||  | | |||  || | |
Db       1802 GTTCAAAGTCTAGCTATTGGTATAATCATCTAATATTATATATATCTCCAGTGCCCCTGA 1861

Qy       1027 TTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACA 1086
              || |||  |  ||  |   |  ||| |     |||  | ||  |     |||   | |
Db       1862 ATTTTATGTTTGATGACTATATATTTGGGCATATATCTTGTTGGATTAGAATAAATAAAA 1921

Qy       1087 ACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                ||  | |  | |     || | ||||||||||||||||||||||||
Db       1922 CACTTTATATTTTCATGAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAA 1970


RESULT 3
US-10-115-123-30
; Sequence 30, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029G30AP1D2
;  CURRENT APPLICATION NUMBER: US/10/115,123
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: PCT/US99/13418
;  PRIOR FILING DATE: 1999-06-15
;  PRIOR APPLICATION NUMBER: 60/089,507
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089,508
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089,509
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089,510
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/090,112
;  PRIOR FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: 60/090,113
;  PRIOR FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 1979
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (968)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-115-123-30

  Query Match             3.9%;  Score 44.2;  DB 3;  Length 1979;
  Best Local Similarity   53.8%;  Pred. No. 0.084;
  Matches   91;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        967 GTTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTT 1026
              ||| ||| || ||   |||   | ||||| |||   ||  | | |||  || | |
Db       1802 GTTCAAAGTCTAGCTATTGGTATAATCATCTAATATTATATATATCTCCAGTGCCCCTGA 1861

Qy       1027 TTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACA 1086
              || |||  |  ||  |   |  ||| |     |||  | ||  |     |||   | |
Db       1862 ATTTTATGTTTGATGACTATATATTTGGGCATATATCTTGTTGGATTAGAATAAATAAAA 1921

Qy       1087 ACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                ||  | |  | |     || | ||||||||||||||||||||||||
Db       1922 CACTTTATATTTTCATGAACTCTAAAAAAAAAAAAAAAAAAAAAAAAAA 1970


RESULT 4
US-08-330-108-16
; Sequence 16, Application US/08330108
; Patent No. 5795752
; GENERAL INFORMATION:
;   APPLICANT:  Smith, Kendall A.
;   TITLE OF INVENTION:  IL-2-Stimulated Gene
;   TITLE OF INVENTION:  Expression
;   NUMBER OF SEQUENCES:  20
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Lahive & Cockfield
;     STREET:  60 State Street
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  U.S.A.
;     ZIP:  02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  ASCII TEXT
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/330,108
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/104,736
;     FILING DATE:
;     APPLICATION NUMBER:  US/07/796,066
;     FILING DATE:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  single stranded
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  no
;     ANTI-SENSE:  no
;     ORIGINAL SOURCE:
;       ORGANISM:  human
;       CELL TYPE:  T-cell blast
;     IMMEDIATE SOURCE:
;       LIBRARY:
;       CLONE:  8D4-T3
US-08-330-108-16

  Query Match             3.8%;  Score 43;  DB 2;  Length 198;
  Best Local Similarity   60.9%;  Pred. No. 0.064;
  Matches   70;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||| ||||  | | ||    |       || ||| ||  | | | | ||| |||  |||||
Db         63 CTGTTTTTATTTTGGTATTAAGTCGTTGCCTTTATTTGTAAAGCTGTTATAAATATATAT 122

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              | || || | | | |   |       || ||   ||||||||||||||||||||
Db        123 TATATAAATATATTAAAAAGGAAAATGTTTCAGAAAAAAAAAAAAAAAAAAAAAA 177


RESULT 5
PCT-US92-10087-16
; Sequence 16, Application PC/TUS9210087
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A.
;    TITLE OF INVENTION:  IL-2-Stimulated Gene
;    TITLE OF INVENTION:  Expression
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lahive & Cockfield
;      STREET:  60 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII TEXT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10087
;      FILING DATE:  19921118
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/796,066
;      FILING DATE:  20-NOVEMBER-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Giulio A. Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  DCI-028PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-227-7400
;      TELEFAX:  617-227-5941
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  198 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  human
;      CELL TYPE:  T-cell blast
;    IMMEDIATE SOURCE:
;      LIBRARY:
;      CLONE:  8D4-T3
PCT-US92-10087-16

  Query Match             3.8%;  Score 43;  DB 6;  Length 198;
  Best Local Similarity   60.9%;  Pred. No. 0.064;
  Matches   70;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy       1021 CTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATAT 1080
              ||| ||||  | | ||    |       || ||| ||  | | | | ||| |||  |||||
Db         63 CTGTTTTTATTTTGGTATTAAGTCGTTGCCTTTATTTGTAAAGCTGTTATAAATATATAT 122

Qy       1081 TGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              | || || | | | |   |       || ||   ||||||||||||||||||||
Db        123 TATATAAATATATTAAAAAGGAAAATGTTTCAGAAAAAAAAAAAAAAAAAAAAAA 177


RESULT 6
US-09-270-767-11283/c
; Sequence 11283, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11283
;   LENGTH: 1637
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-11283

  Query Match             3.8%;  Score 42.8;  DB 3;  Length 1637;
  Best Local Similarity   57.5%;  Pred. No. 0.19;
  Matches   77;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy        997 TAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTT 1056
              | ||||   |      ||| |    | |  |||      ||| |||||      ||||||
Db        135 TCAACAACTGAGAACATTTCGCATTTAGCATTTTCGCCTTCATTATCAATTCAATTTGTT 76

Qy       1057 GCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAA 1116
              ||||  |||  |||   ||||   | ||    |  ||||| ||  |  |  ||||||
Db         75 AACTATGATTTTGAAAAAAAATATATAAAAATATATATAGTCATCTAATGACCAACAAAA 16

Qy       1117 AAAAAAAAAAAAAA 1130
              ||||||||||||||
Db         15 AAAAAAAAAAAAAA 2


RESULT 7
US-08-924-747-25
; Sequence 25, Application US/08924747
; Patent No. 6063570
;  GENERAL INFORMATION:
;    APPLICANT:  MCGONIGLE, BRIAN
;    APPLICANT:  O'KEEFE, DANIEL
;    TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;    TITLE OF INVENTION:  ENZYMES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
```

```
;      SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/924,747
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FLOYD, LINDA AXAMETHY
;      REGISTRATION NUMBER:  33,692
;      REFERENCE/DOCKET NUMBER:  CL-1108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-8112
;      TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  991 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      TISSUE TYPE:  SOYBEAN
;    IMMEDIATE SOURCE:
;      CLONE:  SSM.PK0067.G5
US-08-924-747-25

  Query Match             3.7%;  Score 42.4;  DB 3;  Length 991;
  Best Local Similarity   54.5%;  Pred. No. 0.19;
  Matches   85;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        980 AGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAA 1039
              ||| |||  || |   | ||  || |  || |  | ||     | || |||| || | ||
Db        800 AGTATGATTTTGTTGGGAAACAATTATCTTGTTGTGAGCAAAGGATTGTTCTGTTTTAAA 859

Qy       1040 TATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCA 1099
              | | ||   || |  ||   |  ||| ||    | ||  ||| ||        |
Db        860 TTTAATTGACTGTGATTTGGTTGGGTATTGGCTATTTTAATTTTAACTAAAAAAAGTGTT 919

Qy       1100 CCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              |   |||    | ||||||||||||||||||||||
Db        920 CAGTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 955


RESULT 8
US-09-247-373B-25
; Sequence 25, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
;  APPLICANT:  MCGONIGLE, BRIAN
;  APPLICANT:  O'KEEFE, DANIEL
;  TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE:  CL-1108-A
;  CURRENT APPLICATION NUMBER: US/09/247,373B
;  CURRENT FILING DATE:  1999-02-10
;  PRIOR APPLICATION NUMBER:  08/924,747
;  PRIOR FILING DATE:  1997-09-05
;  NUMBER OF SEQ ID NOS:  56
;  SOFTWARE:  Microsoft Office 97
; SEQ ID NO 25
;   LENGTH: 991
;   TYPE: DNA
;   ORGANISM: SOYBEAN
US-09-247-373B-25

  Query Match             3.7%;  Score 42.4;  DB 3;  Length 991;
  Best Local Similarity   54.5%;  Pred. No. 0.19;
  Matches   85;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        980 AGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAA 1039
              ||| |||  || |   | ||  || |  || |  | ||     | || |||| || | ||
Db        800 AGTATGATTTTGTTGGGAAACAATTATCTTGTTGTGAGCAAAGGATTGTTCTGTTTTAAA 859

Qy       1040 TATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCA 1099
              | | ||   || |  ||   |  ||| ||    | ||  ||| ||        |
Db        860 TTTAATTGACTGTGATTTGGTTGGGTATTGGCTATTTTAATTTTAACTAAAAAAAGTGTT 919

Qy       1100 CCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              |   |||    | ||||||||||||||||||||||
Db        920 CAGTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 955


RESULT 9
US-09-296-715-25
; Sequence 25, Application US/09296715
; Patent No. 6171839
; GENERAL INFORMATION:
;   APPLICANT:  MCGONIGLE, BRIAN
;   APPLICANT:  O'KEEFE, DANIEL
;   TITLE OF INVENTION:  SOYBEAN GLUTATHIONE-S-TRANSFERASE
;   TITLE OF INVENTION:  ENZYMES
;   NUMBER OF SEQUENCES:  32
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  E.I. DU PONT DE NEMOURS AND COMPANY
;     STREET:  1007 MARKET STREET
;     CITY:  WILMINGTON
;     STATE:  DELAWARE
;     COUNTRY:  UNITED STATES OF AMERICA
;     ZIP:  19898
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  DISKETTE, 3.50 INCH
;     COMPUTER:  IBM PC COMPATIBLE
;     OPERATING SYSTEM:  MICROSOFT WORD FOR WINDOWS 95
;     SOFTWARE:  MICROSOFT WORD VERSION 7.0A
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/296,715
;     FILING DATE:
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  FLOYD, LINDA AXAMETHY
;     REGISTRATION NUMBER:  33,692
;     REFERENCE/DOCKET NUMBER:  CL-1108
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  302-892-8112
;     TELEFAX:  302-773-0164
;  INFORMATION FOR SEQ ID NO:  25:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  991 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   ORIGINAL SOURCE:
;     TISSUE TYPE:  SOYBEAN
;   IMMEDIATE SOURCE:
;     CLONE:  SSM.PK0067.G5
US-09-296-715-25

  Query Match             3.7%;  Score 42.4;  DB 3;  Length 991;
  Best Local Similarity   54.5%;  Pred. No. 0.19;
  Matches   85;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        980 AGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAA 1039
              ||| |||  || |   | ||  || |  || |  | ||     | || |||| || | ||
Db        800 AGTATGATTTTGTTGGGAAACAATTATCTTGTTGTGAGCAAAGGATTGTTCTGTTTTAAA 859

Qy       1040 TATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCA 1099
              | | ||   || |  ||   |  ||| ||    | ||  ||| ||        |
Db        860 TTTAATTGACTGTGATTTGGTTGGGTATTGGCTATTTTAATTTTAACTAAAAAAAGTGTT 919

Qy       1100 CCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              |   |||    | ||||||||||||||||||||||
Db        920 CAGTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 955
```

```
RESULT 10
US-09-832-129-13
; Sequence 13, Application US/09832129
; Patent No. 6936691
; GENERAL INFORMATION:
;  APPLICANT: Fiscella et al.
;  TITLE OF INVENTION: 19 Human secreted proteins
;  FILE REFERENCE: PZ045P1
;  CURRENT APPLICATION NUMBER: US/09/832,129
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: PCT/US00/28664
;  PRIOR FILING DATE: 2000-10-17
;  PRIOR APPLICATION NUMBER: 60/163,085
;  PRIOR FILING DATE: 1999-11-02
;  PRIOR APPLICATION NUMBER: 60/172,411
;  PRIOR FILING DATE: 1999-12-17
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 3265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-832-129-13

  Query Match             3.7%;  Score 42;  DB 3;  Length 3265;
  Best Local Similarity   59.0%;  Pred. No. 0.42;
  Matches   72;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy       1014 TTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAA 1073
              |||    ||  ||||  | || | |  || | |||||||||| |  ||     ||  |||
Db       3131 TTAAACTCTTCTTTTCTTTTTATTATAATGACCAGCTTTTGGTATTTCATTGTTACCAAG 3190

Qy       1074 CTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAA 1133
              |    || || ||      |   || |||  |     | ||||||||||||||||||||
Db       3191 TTCTATTTTTAGAATAAAATTGTTCTCCTTCTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3250

Qy       1134 AA 1135
              ||
Db       3251 AA 3252


RESULT 11
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             3.7%;  Score 41.6;  DB 3;  Length 640681;
  Best Local Similarity   57.8%;  Pred. No. 5.5;
  Matches   74;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Qy        968 TTGAAATTCAAGAGTTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTT 1027
              | || |   | |  ||||   || | || |||| |||| |||||   ||     || |||
Db     527360 TAGATAAAAATGTTTTTGTTTTTTTGATCTAAAAATAATTTTTTTGATAACATTTGATTT 527301

Qy       1028 TTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAA 1087
              | || ||||| || || | | ||| |||  ||   |  | ||| ||| |   || ||
Db     527300 ATTCATAGTCAAAATGATAATCATTTATTGTATAAAATAAAAAAAATTTAAAGATTATAA 527241

Qy       1088 CTCTTTAA 1095
              |  | ||
Db     527240 TTGATAAA 527233


RESULT 12
US-10-033-301-23
; Sequence 23, Application US/10033301
; Patent No. 6930172
; GENERAL INFORMATION:
;  APPLICANT: Botstein,David
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Fong,Sherman
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Roy,Margaret Ann
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2930R1C6
;  CURRENT APPLICATION NUMBER: US/10/033,301
;  CURRENT FILING DATE:  2001-12-27
;  PRIOR APPLICATION NUMBER: 60/095,325
;  PRIOR FILING DATE: 1998-08-04
;  PRIOR APPLICATION NUMBER: 60/112,851
;  PRIOR FILING DATE: 1998-12-16
;  PRIOR APPLICATION NUMBER: 60/113,145
;  PRIOR FILING DATE: 1998-12-16
;  PRIOR APPLICATION NUMBER: 60/113,511
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: 60/115,558
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/115,565
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/115,733
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/119,341
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/119,537
;  PRIOR FILING DATE: 1999-02-10
;  PRIOR APPLICATION NUMBER: 60/119,965
;  PRIOR FILING DATE: 1999-02-12
;  PRIOR APPLICATION NUMBER: 60/162,506
;  PRIOR FILING DATE: 1999-10-29
;  PRIOR APPLICATION NUMBER: 60/170,262
;  PRIOR FILING DATE: 1999-12-09
;  PRIOR APPLICATION NUMBER: 60/187,202
;  PRIOR FILING DATE: 2000-03-03
;  PRIOR APPLICATION NUMBER: PCT/US99/12252
;  PRIOR FILING DATE: 1999-06-02
;  PRIOR APPLICATION NUMBER: PCT/US99/28634
;  PRIOR FILING DATE: 1999-12-01
;  PRIOR APPLICATION NUMBER: PCT/US99/28551
;  PRIOR FILING DATE: 1999-12-02
;  PRIOR APPLICATION NUMBER: PCT/US00/03565
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: PCT/US00/04414
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: PCT/US00/05841
;  PRIOR FILING DATE: 2000 -03-02
;  PRIOR APPLICATION NUMBER: PCT/US00/08439
;  PRIOR FILING DATE: 2000-03-30
;  PRIOR APPLICATION NUMBER: PCT/US00/14941
```

```
;  PRIOR FILING DATE: 2000-05-30
;  PRIOR APPLICATION NUMBER: PCT/US00/15264
;  PRIOR FILING DATE: 2000-06-02
;  PRIOR APPLICATION NUMBER: PCT/US00/32678
;  PRIOR FILING DATE: 2000-12-01
;  NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 23
;   LENGTH: 2186
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-033-301-23

  Query Match             3.6%;  Score 40.4;  DB 3;  Length 2186;
  Best Local Similarity   53.9%;  Pred. No. 0.96;
  Matches   83;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        982 TTTGAGGTTATCATGTAAACATAAGTTTTTCTTTAGGGGCTGGTTTTTCTATTGTCAATA 1041
              | ||  |  | ||   ||   |  |  | ||||| | ||||  |||| |||| | |
Db       2001 TCTGGCGCGAACACTGGAATGCATATACTACTTTATGTGCTGTGTTTTTTATTCTTGGAT 2060

Qy       1042 TCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCTTTAAGTCACC 1101
              |||  | |||| |  | || |     | | |||| |||   |       ||     |
Db       2061 ACATTTGATTTTTTCACGTAAGTCCACATATACTTCTATAAGAGCGTGACTTGTAATAAA 2120

Qy       1102 TCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                 ||  | |  ||||||||||||||||||||
Db       2121 GGGTTAATGAAGAAAAAAAAAAAAAAAAAAAAA 2154


RESULT 13
US-09-008-697A-11
; Sequence 11, Application US/09008697A
; Patent No. 6197504
;  GENERAL INFORMATION:
;    APPLICANT:  Chow, King Lau
;    TITLE OF INVENTION:  USES OF MAB-21
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/008,697A
;      FILING DATE:  January 19, 1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Chan, Albert Wai-Kit
;      REGISTRATION NUMBER:  36,479
;      REFERENCE/DOCKET NUMBER:  50752-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-0400
;      TELEFAX:  (212) 391-0525
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1278 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-008-697A-11

  Query Match             3.5%;  Score 40;  DB 3;  Length 1278;
  Best Local Similarity   59.8%;  Pred. No. 0.97;
  Matches   67;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy       1024 GTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGT 1083
              |||| |   || | | |||||    | |       |||       | |   ||| |||  ||
Db       1166 GTTTGTAATTTTTGATTATCAATTCCCTCAAACTCTTTCTCACTCAACTACTGATACAGT 1225

Qy       1084 ACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
                ||    | ||| | |  | ||| ||  ||||||||||||||||||||||||
Db       1226 TCATACTTATAAATAAATTATTTTCTCCCAAAAAAAAAAAAAAAAAAAAAAA 1277


RESULT 14
US-08-148-209A-1
; Sequence 1, Application US/08148209A
; Patent No. 5556780
;  GENERAL INFORMATION:
;    APPLICANT:  Dzau, Victor J
;    APPLICANT:  Mukoyama, Masashi
;    TITLE OF INVENTION:  TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/148,209A
;      FILING DATE:  05-NOV-1993
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rowland, Bertram I
;      REGISTRATION NUMBER:  20,015
;      REFERENCE/DOCKET NUMBER:  A-58491-1/BIR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2862 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  132..1223
US-08-148-209A-1

  Query Match             3.5%;  Score 40;  DB 2;  Length 2862;
  Best Local Similarity   61.5%;  Pred. No. 1.4;
  Matches   64;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       1032 ATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTACAAACTTATATTGTACAACTCT 1091
              || |  ||     | | |  || |  |||   ||||||   | || ||    | | |||
Db       2747 ATAGCTAAATATGTAATCAGTTATGACTTTGTGTTTTAAGCAATTTTACACAAAATCTCG 2806

Qy       1092 TTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAAAAAAA 1135
              | || | |  || || ||   ||||||||||||||||||||||
Db       2807 TAAAATAAAATCATTACTGGGAAAAAAAAAAAAAAAAAAAAAAA 2850


RESULT 15
US-09-881-556A-1
; Sequence 1, Application US/09881556A
; Patent No. 6808900
; GENERAL INFORMATION:
;  APPLICANT: Simonsen, J. Neil
```

; TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereto and Diagnosti
; TITLE OF INVENTION: Therapeutic Compositions Thereof
; FILE REFERENCE: SASK-008/01US
; CURRENT APPLICATION NUMBER: US/09/881,556A
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/212,083
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-881-556A-1

Query Match 3.5%; Score 39.6; DB 3; Length 1380;
Best Local Similarity 57.1%; Pred. No. 1.3;
Matches 72; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
Qy       1010 TTCTTTAGGGGCTGGTTTTTCTATTGTCAATATCATCAGCTTTTGTTGCTTATGGTTTTA 1069
              ||  | ||||||  |  ||   |||| |  || | |  |  ||    |  ||  || ||
Db       1239 TTAATAAGGGGTAGAATTAGATATTTTTCATGTAAATAAATTAGCGTTATTGAGGATTAT 1298

Qy       1070 CAAACTTATATTGTACAACTCTTTAAGTCACCTCTTTGCTTACAAAAAAAAAAAAAAAAAA 1129
                 |  || ||  || |  | || |  | |  | ||   | | ||||||||||||||||||
Db       1299 TCGAAATAAATAATAGAGATATTAAGTTTAGTTTTTATTTAAAAAAAAAAAAAAAAAAAAA 1358

Qy       1130 AAAAAA 1135
              ||||||
Db       1359 AAAAAA 1364
```

Search completed: March 20, 2006, 14:05:56
Job time : 229.218 secs

THIS PAGE LEFT BLANK